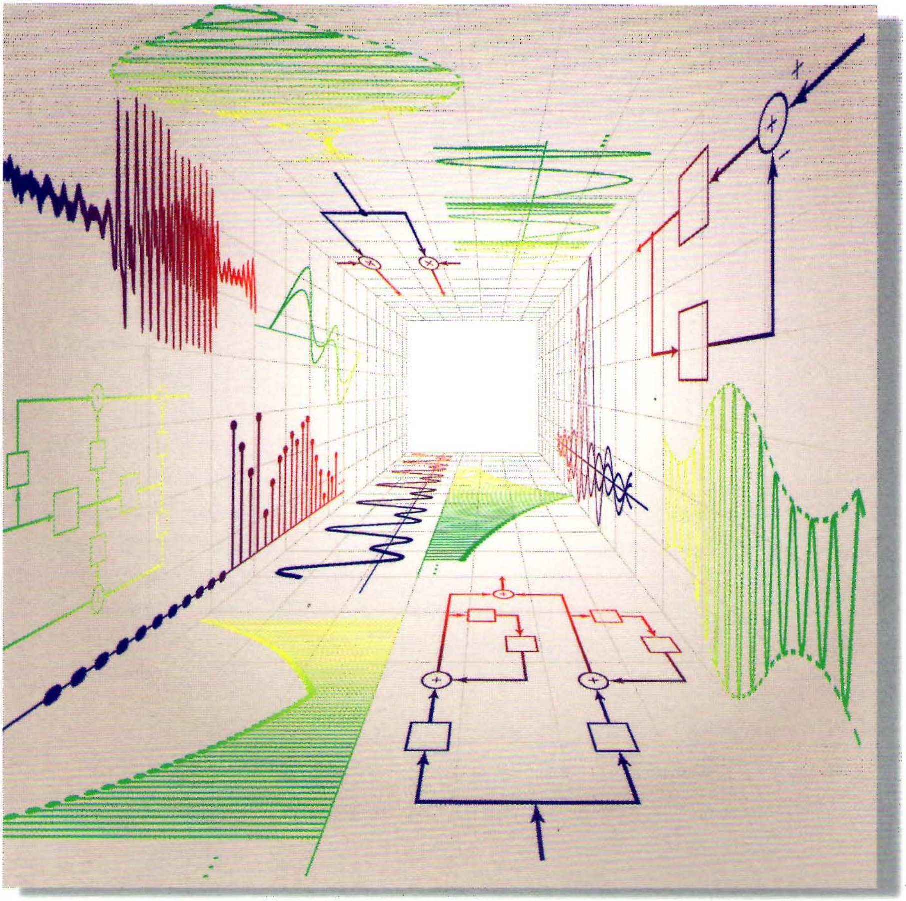
INTERNATIONAL EDITION

SIGNALS & SYSTEMS

SECOND EDITION



ALAN V. OPPENHEIM
ALAN S. WILLSKY
WITH S. HAMID NAWAB

PRENTICE HALL SIGNAL PROCESSING SERIES
ALAN V. OPPENHEIM, SERIES EDITOR

SECOND EDITION

SIGNALS & SYSTEMS

ALAN V. OPPENHEIM
ALAN S. WILLSKY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

WITH

S. HAMID NAWAB
BOSTON UNIVERSITY

Prentice Hall

PRENTICE-HALL INTERNATIONAL, INC.

CONTENTS

PREFACE XVII

ACKNOWLEDGMENTS XXV

FOREWORD XXVII

1 SIGNALS AND SYSTEMS 1

1.0 Introduction 1

1.1 Continuous-Time and Discrete-Time Signals 1
- 1.1.1 Examples and Mathematical Representation 1
- 1.1.2 Signal Energy and Power 5

1.2 Transformations of the Independent Variable 7
- 1.2.1 Examples of Transformations of the Independent Variable 8
- 1.2.2 Periodic Signals 11
- 1.2.3 Even and Odd Signals 13

1.3 Exponential and Sinusoidal Signals 14
- 1.3.1 Continuous-Time Complex Exponential and Sinusoidal Signals 15
- 1.3.2 Discrete-Time Complex Exponential and Sinusoidal Signals 21
- 1.3.3 Periodicity Properties of Discrete-Time Complex Exponentials 25

1.4 The Unit Impulse and Unit Step Functions 30
- 1.4.1 The Discrete-Time Unit Impulse and Unit Step Sequences 30
- 1.4.2 The Continuous-Time Unit Step and Unit Impulse Functions 32

1.5 Continuous-Time and Discrete-Time Systems 38
- 1.5.1 Simple Examples of Systems 39
- 1.5.2 Interconnections of Systems 41

1.6 Basic System Properties 44
- 1.6.1 Systems with and without Memory 44
- 1.6.2 Invertibility and Inverse Systems 45
- 1.6.3 Causality 46
- 1.6.4 Stability 48
- 1.6.5 Time Invariance 50
- 1.6.6 Linearity 53

1.7 Summary 56

Problems 57

2 LINEAR TIME-INVARIANT SYSTEMS 74

2.0 Introduction 74

2.1 Discrete-Time LTI Systems: The Convolution Sum 75

2.1.1 The Representation of Discrete-Time Signals in Terms of Impulses 75
2.1.2 The Discrete-Time Unit Impulse Response and the Convolution-Sum Representation of LTI Systems 77

2.2 Continuous-Time LTI Systems: The Convolution Integral 90
2.2.1 The Representation of Continuous-Time Signals in Terms of Impulses 90
2.2.2 The Continuous-Time Unit Impulse Response and the Convolution Integral Representation of LTI Systems 94

2.3 Properties of Linear Time-Invariant Systems 103
2.3.1 The Commutative Property 104
2.3.2 The Distributive Property 104
2.3.3 The Associative Property 107
2.3.4 LTI Systems with and without Memory 108
2.3.5 Invertibility of LTI Systems 109
2.3.6 Causality for LTI Systems 112
2.3.7 Stability for LTI Systems 113
2.3.8 The Unit Step Response of an LTI System 115

2.4 Causal LTI Systems Described by Differential and Difference Equations 116
2.4.1 Linear Constant-Coefficient Differential Equations 117
2.4.2 Linear Constant-Coefficient Difference Equations 121
2.4.3 Block Diagram Representations of First-Order Systems Described by Differential and Difference Equations 124

2.5 Singularity Functions 127
2.5.1 The Unit Impulse as an Idealized Short Pulse 128
2.5.2 Defining the Unit Impulse through Convolution 131
2.5.3 Unit Doublets and Other Singularity Functions 132

2.6 Summary 137

Problems 137

3 FOURIER SERIES REPRESENTATION OF PERIODIC SIGNALS 177

3.0 Introduction 177

3.1 A Historical Perspective 178

3.2 The Response of LTI Systems to Complex Exponentials 182

3.3 Fourier Series Representation of Continuous-Time Periodic Signals 186
3.3.1 Linear Combinations of Harmonically Related Complex Exponentials 186
3.3.2 Determination of the Fourier Series Representation of a Continuous-Time Periodic Signal 190

3.4 Convergence of the Fourier Series 195

3.5 Properties of Continuous-Time Fourier Series 202
  3.5.1 Linearity 202
  3.5.2 Time Shifting 202
  3.5.3 Time Reversal 203
  3.5.4 Time Scaling 204
  3.5.5 Multiplication 204
  3.5.6 Conjugation and Conjugate Symmetry 204
  3.5.7 Parseval's Relation for Continuous-Time Periodic Signals 205
  3.5.8 Summary of Properties of the Continuous-Time Fourier Series 205
  3.5.9 Examples 205
3.6 Fourier Series Representation of Discrete-Time Periodic Signals 211
  3.6.1 Linear Combinations of Harmonically Related Complex Exponentials 211
  3.6.2 Determination of the Fourier Series Representation of a Periodic Signal 212
3.7 Properties of Discrete-Time Fourier Series 221
  3.7.1 Multiplication 222
  3.7.2 First Difference 222
  3.7.3 Parseval's Relation for Discrete-Time Periodic Signals 223
  3.7.4 Examples 223
3.8 Fourier Series and LTI Systems 226
3.9 Filtering 231
  3.9.1 Frequency-Shaping Filters 232
  3.9.2 Frequency-Selective Filters 236
3.10 Examples of Continuous-Time Filters Described by Differential Equations 239
  3.10.1 A Simple *RC* Lowpass Filter 239
  3.10.2 A Simple *RC* Highpass Filter 241
3.11 Examples of Discrete-Time Filters Described by Difference Equations 244
  3.11.1 First-Order Recursive Discrete-Time Filters 244
  3.11.2 Nonrecursive Discrete-Time Filters 245
3.12 Summary 249
Problems 250

4 THE CONTINUOUS-TIME FOURIER TRANSFORM 284

4.0 Introduction 284
4.1 Representation of Aperiodic Signals: The Continuous-Time Fourier Transform 285
  4.1.1 Development of the Fourier Transform Representation of an Aperiodic Signal 285
  4.1.2 Convergence of Fourier Transforms 289
  4.1.3 Examples of Continuous-Time Fourier Transforms 290

4.2 The Fourier Transform for Periodic Signals 296
4.3 Properties of the Continuous-Time Fourier Transform 300
  4.3.1 Linearity 301
  4.3.2 Time Shifting 301
  4.3.3 Conjugation and Conjugate Symmetry 303
  4.3.4 Differentiation and Integration 306
  4.3.5 Time and Frequency Scaling 308
  4.3.6 Duality 309
  4.3.7 Parseval's Relation 312
4.4 The Convolution Property 314
  4.4.1 Examples 317
4.5 The Multiplication Property 322
  4.5.1 Frequency-Selective Filtering with Variable Center Frequency 325
4.6 Tables of Fourier Properties and of Basic Fourier Transform Pairs 328
4.7 Systems Characterized by Linear Constant-Coefficient Differential Equations 330
4.8 Summary 333
Problems 334

5 THE DISCRETE-TIME FOURIER TRANSFORM 358

5.0 Introduction 358
5.1 Representation of Aperiodic Signals: The Discrete-Time Fourier Transform 359
  5.1.1 Development of the Discrete-Time Fourier Transform 359
  5.1.2 Examples of Discrete-Time Fourier Transforms 362
  5.1.3 Convergence Issues Associated with the Discrete-Time Fourier Transform 366
5.2 The Fourier Transform for Periodic Signals 367
5.3 Properties of the Discrete-Time Fourier Transform 372
  5.3.1 Periodicity of the Discrete-Time Fourier Transform 373
  5.3.2 Linearity of the Fourier Transform 373
  5.3.3 Time Shifting and Frequency Shifting 373
  5.3.4 Conjugation and Conjugate Symmetry 375
  5.3.5 Differencing and Accumulation 375
  5.3.6 Time Reversal 376
  5.3.7 Time Expansion 377
  5.3.8 Differentiation in Frequency 380
  5.3.9 Parseval's Relation 380
5.4 The Convolution Property 382
  5.4.1 Examples 383
5.5 The Multiplication Property 388
5.6 Tables of Fourier Transform Properties and Basic Fourier Transform Pairs 390

5.7 Duality 390
  5.7.1 Duality in the Discrete-Time Fourier Series 391
  5.7.2 Duality between the Discrete-Time Fourier Transform and the Continuous-Time Fourier Series 395
5.8 Systems Characterized by Linear Constant-Coefficient Difference Equations 396
5.9 Summary 399
Problems 400

6 TIME AND FREQUENCY CHARACTERIZATION OF SIGNALS AND SYSTEMS 423

6.0 Introduction 423
6.1 The Magnitude-Phase Representation of the Fourier Transform 423
6.2 The Magnitude-Phase Representation of the Frequency Response of LTI Systems 427
  6.2.1 Linear and Nonlinear Phase 428
  6.2.2 Group Delay 430
  6.2.3 Log-Magnitude and Bode Plots 436
6.3 Time-Domain Properties of Ideal Frequency-Selective Filters 439
6.4 Time-Domain and Frequency-Domain Aspects of Nonideal Filters 444
6.5 First-Order and Second-Order Continuous-Time Systems 448
  6.5.1 First-Order Continuous-Time Systems 448
  6.5.2 Second-Order Continuous-Time Systems 451
  6.5.3 Bode Plots for Rational Frequency Responses 456
6.6 First-Order and Second-Order Discrete-Time Systems 461
  6.6.1 First-Order Discrete-Time Systems 461
  6.6.2 Second-Order Discrete-Time Systems 465
6.7 Examples of Time- and Frequency-Domain Analysis of Systems 472
  6.7.1 Analysis of an Automobile Suspension System 473
  6.7.2 Examples of Discrete-Time Nonrecursive Filters 476
6.8 Summary 482
Problems 483

7 SAMPLING 514

7.0 Introduction 514
7.1 Representation of a Continuous-Time Signal by Its Samples: The Sampling Theorem 515
  7.1.1 Impulse-Train Sampling 516
  7.1.2 Sampling with a Zero-Order Hold 520

7.2 Reconstruction of a Signal from Its Samples Using Interpolation 522
7.3 The Effect of Undersampling: Aliasing 527
7.4 Discrete-Time Processing of Continuous-Time Signals 534
  7.4.1 Digital Differentiator 541
  7.4.2 Half-Sample Delay 543
7.5 Sampling of Discrete-Time Signals 545
  7.5.1 Impulse-Train Sampling 545
  7.5.2 Discrete-Time Decimation and Interpolation 549
7.6 Summary 555
Problems 556

8 COMMUNICATION SYSTEMS 582

8.0 Introduction 582
8.1 Complex Exponential and Sinusoidal Amplitude Modulation 583
  8.1.1 Amplitude Modulation with a Complex Exponential Carrier 583
  8.1.2 Amplitude Modulation with a Sinusoidal Carrier 585
8.2 Demodulation for Sinusoidal AM 587
  8.2.1 Synchronous Demodulation 587
  8.2.2 Asynchronous Demodulation 590
8.3 Frequency-Division Multiplexing 594
8.4 Single-Sideband Sinusoidal Amplitude Modulation 597
8.5 Amplitude Modulation with a Pulse-Train Carrier 601
  8.5.1 Modulation of a Pulse-Train Carrier 601
  8.5.2 Time-Division Multiplexing 604
8.6 Pulse-Amplitude Modulation 604
  8.6.1 Pulse-Amplitude Modulated Signals 604
  8.6.2 Intersymbol Interference in PAM Systems 607
  8.6.3 Digital Pulse-Amplitude and Pulse-Code Modulation 610
8.7 Sinusoidal Frequency Modulation 611
  8.7.1 Narrowband Frequency Modulation 613
  8.7.2 Wideband Frequency Modulation 615
  8.7.3 Periodic Square-Wave Modulating Signal 617
8.8 Discrete-Time Modulation 619
  8.8.1 Discrete-Time Sinusoidal Amplitude Modulation 619
  8.8.2 Discrete-Time Transmodulation 623
8.9 Summary 623
Problems 625

9 THE LAPLACE TRANSFORM 654

9.0 Introduction 654
9.1 The Laplace Transform 655
9.2 The Region of Convergence for Laplace Transforms 662

9.3 The Inverse Laplace Transform 670
9.4 Geometric Evaluation of the Fourier Transform from the Pole-Zero Plot 674
  9.4.1 First-Order Systems 676
  9.4.2 Second-Order Systems 677
  9.4.3 All-Pass Systems 681
9.5 Properties of the Laplace Transform 682
  9.5.1 Linearity of the Laplace Transform 683
  9.5.2 Time Shifting 684
  9.5.3 Shifting in the s-Domain 685
  9.5.4 Time Scaling 685
  9.5.5 Conjugation 687
  9.5.6 Convolution Property 687
  9.5.7 Differentiation in the Time Domain 688
  9.5.8 Differentiation in the s-Domain 688
  9.5.9 Integration in the Time Domain 690
  9.5.10 The Initial- and Final-Value Theorems 690
  9.5.11 Table of Properties 691
9.6 Some Laplace Transform Pairs 692
9.7 Analysis and Characterization of LTI Systems Using the Laplace Transform 693
  9.7.1 Causality 693
  9.7.2 Stability 695
  9.7.3 LTI Systems Characterized by Linear Constant-Coefficient Differential Equations 698
  9.7.4 Examples Relating System Behavior to the System Function 701
  9.7.5 Butterworth Filters 703
9.8 System Function Algebra and Block Diagram Representations 706
  9.8.1 System Functions for Interconnections of LTI Systems 707
  9.8.2 Block Diagram Representations for Causal LTI Systems Described by Differential Equations and Rational System Functions 708
9.9 The Unilateral Laplace Transform 714
  9.9.1 Examples of Unilateral Laplace Transforms 714
  9.9.2 Properties of the Unilateral Laplace Transform 716
  9.9.3 Solving Differential Equations Using the Unilateral Laplace Transform 719
9.10 Summary 720
Problems 721

10 THE z-TRANSFORM 741

10.0 Introduction 741
10.1 The z-Transform 741
10.2 The Region of Convergence for the z-Transform 748

10.3 The Inverse z-Transform 757

10.4 Geometric Evaluation of the Fourier Transform from the Pole-Zero Plot 763

- 10.4.1 First-Order Systems 763
- 10.4.2 Second-Order Systems 765

10.5 Properties of the z-Transform 767

- 10.5.1 Linearity 767
- 10.5.2 Time Shifting 767
- 10.5.3 Scaling in the z-Domain 768
- 10.5.4 Time Reversal 769
- 10.5.5 Time Expansion 769
- 10.5.6 Conjugation 770
- 10.5.7 The Convolution Property 770
- 10.5.8 Differentiation in the z-Domain 772
- 10.5.9 The Initial-Value Theorem 773
- 10.5.10 Summary of Properties 774

10.6 Some Common z-Transform Pairs 774

10.7 Analysis and Characterization of LTI Systems Using z-Transforms 774

- 10.7.1 Causality 776
- 10.7.2 Stability 777
- 10.7.3 LTI Systems Characterized by Linear Constant-Coefficient Difference Equations 779
- 10.7.4 Examples Relating System Behavior to the System Function 781

10.8 System Function Algebra and Block Diagram Representations 783

- 10.8.1 System Functions for Interconnections of LTI Systems 784
- 10.8.2 Block Diagram Representations for Causal LTI Systems Described by Difference Equations and Rational System Functions 784

10.9 The Unilateral z-Transform 789

- 10.9.1 Examples of Unilateral z-Transforms and Inverse Transforms 790
- 10.9.2 Properties of the Unilateral z-Transform 792
- 10.9.3 Solving Difference Equations Using the Unilateral z-Transform 795

10.10 Summary 796

Problems 797

11 LINEAR FEEDBACK SYSTEMS 816

11.0 Introduction 816

11.1 Linear Feedback Systems 819

11.2 Some Applications and Consequences of Feedback 820

- 11.2.1 Inverse System Design 820
- 11.2.2 Compensation for Nonideal Elements 821
- 11.2.3 Stabilization of Unstable Systems 823

PREFACE

This book is the second edition of a text designed for undergraduate courses in signals and systems. While such courses are frequently found in electrical engineering curricula, the concepts and techniques that form the core of the subject are of fundamental importance in all engineering disciplines. In fact, the scope of potential and actual applications of the methods of signal and system analysis continues to expand as engineers are confronted with new challenges involving the synthesis or analysis of complex processes. For these reasons we feel that a course in signals and systems not only is an essential element in an engineering program but also can be one of the most rewarding, exciting, and useful courses that engineering students take during their undergraduate education.

Our treatment of the subject of signals and systems in this second edition maintains the same general philosophy as in the first edition but with significant rewriting, restructuring, and additions. These changes are designed to help both the instructor in presenting the subject material and the student in mastering it. In the preface to the first edition we stated that our overall approach to signals and systems had been guided by the continuing developments in technologies for signal and system design and implementation, which made it increasingly important for a student to have equal familiarity with techniques suitable for analyzing and synthesizing both continuous-time and discrete-time systems. As we write the preface to this second edition, that observation and guiding principle are even more true than before. Thus, while students studying signals and systems should certainly have a solid foundation in disciplines based on the laws of physics, they must also have a firm grounding in the use of computers for the analysis of phenomena and the implementation of systems and algorithms. As a consequence, engineering curricula now reflect a blend of subjects, some involving continuous-time models and others focusing on the use of computers and discrete representations. For these reasons, signals and systems courses that bring discrete-time and continuous-time concepts together in a unified way play an increasingly important role in the education of engineering students and in their preparation for current and future developments in their chosen fields.

It is with these goals in mind that we have structured this book to develop in parallel the methods of analysis for continuous-time and discrete-time signals and systems. This approach also offers a distinct and extremely important pedagogical advantage. Specifically, we are able to draw on the similarities between continuous- and discrete-time methods in order to share insights and intuition developed in each domain. Similarly, we can exploit the differences between them to sharpen an understanding of the distinct properties of each.

In organizing the material both originally and now in the second edition, we have also considered it essential to introduce the student to some of the important uses of the basic methods that are developed in the book. Not only does this provide the student with an appreciation for the range of applications of the techniques being learned and for directions for further study, but it also helps to deepen understanding of the subject. To achieve this goal we include introductory treatments on the subjects of filtering, communications, sampling, discrete-time processing of continuous-time signals, and feedback. In fact, in one of the major changes in this second edition, we have introduced the concept of frequency-domain filtering very early in our treatment of Fourier analysis in order to provide both motivation for and insight into this very important topic. In addition, we have again included an up-to-date bibliography at the end of the book in order to assist the student who is interested in pursuing additional and more advanced studies of the methods and applications of signal and system analysis.

The organization of the book reflects our conviction that full mastery of a subject of this nature cannot be accomplished without a significant amount of practice in using and applying the tools that are developed. Consequently, in the second edition we have significantly increased the number of worked examples within each chapter. We have also enhanced one of the key assets of the first edition, namely the end-of-chapter homework problems. As in the first edition, we have included a substantial number of problems, totaling more than 600 in number. A majority of the problems included here are new and thus provide additional flexibility for the instructor in preparing homework assignments. In addition, in order to enhance the utility of the problems for both the student and the instructor we have made a number of other changes to the organization and presentation of the problems. In particular, we have organized the problems in each chapter under several specific headings, each of which spans the material in the entire chapter but with a different objective. The first two sections of problems in each chapter emphasize the mechanics of using the basic concepts and methods presented in the chapter. For the first of these two sections, which has the heading Basic Problems with Answers, we have also provided answers (but not solutions) at the end of the book. These answers provide a simple and immediate way for the student to check his or her understanding of the material. The problems in this first section are generally appropriate for inclusion in homework sets. Also, in order to give the instructor additional flexibility in assigning homework problems, we have provided a second section of Basic Problems for which answers have not been included.

A third section of problems in each chapter, organized under the heading of Advanced Problems, is oriented toward exploring and elaborating upon the foundations and practical implications of the material in the text. These problems often involve mathematical derivations and more sophisticated use of the concepts and methods presented in the chapter. Some chapters also include a section of Extension Problems which involve extensions of material presented in the chapter and/or involve the use of knowledge from applications that are outside the scope of the main text (such as advanced circuits or mechanical systems). The overall variety and quantity of problems in each chapter will hopefully provide students with the means to develop their understanding of the material and instructors with considerable flexibility in putting together homework sets that are tailored to the specific needs of their students. A solutions manual is also available to instructors through the publisher.

Another significant additional enhancement to this second edition is the availability of the companion book *Explorations in Signals and Systems Using MATLAB* by Buck, Daniel, and Singer. This book contains MATLAB™-based computer exercises for each topic in the text, and should be of great assistance to both instructor and student.

Students using this book are assumed to have a basic background in calculus as well as some experience in manipulating complex numbers and some exposure to differential equations. With this background, the book is self-contained. In particular, no prior experience with system analysis, convolution, Fourier analysis, or Laplace and z-transforms is assumed. Prior to learning the subject of signals and systems most students will have had a course such as basic circuit theory for electrical engineers or fundamentals of dynamics for mechanical engineers. Such subjects touch on some of the basic ideas that are developed more fully in this text. This background can clearly be of great value to students in providing additional perspective as they proceed through the book.

The Foreword, which follows this preface, is written to offer the reader motivation and perspective for the subject of signals and systems in general and our treatment of it in particular. We begin Chapter 1 by introducing some of the elementary ideas related to the mathematical representation of signals and systems. In particular we discuss transformations (such as time shifts and scaling) of the independent variable of a signal. We also introduce some of the most important and basic continuous-time and discrete-time signals, namely real and complex exponentials and the continuous-time and discrete-time unit step and unit impulse. Chapter 1 also introduces block diagram representations of interconnections of systems and discusses several basic system properties such as causality, linearity and time-invariance. In Chapter 2 we build on these last two properties, together with the sifting property of unit impulses to develop the convolution-sum representation for discrete-time linear, time-invariant (LTI) systems and the convolution integral representation for continuous-time LTI systems. In this treatment we use the intuition gained from our development of the discrete-time case as an aid in deriving and understanding its continuous-time counterpart. We then turn to a discussion of causal, LTI systems characterized by linear constant-coefficient differential and difference equations. In this introductory discussion we review the basic ideas involved in solving linear differential equations (to which most students will have had some previous exposure) and we also provide a discussion of analogous methods for linear difference equations. However, the primary focus of our development in Chapter 2 is not on methods of solution, since more convenient approaches are developed later using transform methods. Instead, in this first look, our intent is to provide the student with some appreciation for these extremely important classes of systems, which will be encountered often in subsequent chapters. Finally, Chapter 2 concludes with a brief discussion of singularity functions—steps, impulses, doublets, and so forth—in the context of their role in the description and analysis of continuous-time LTI systems. In particular, we stress the interpretation of these signals in terms of how they are defined under convolution—that is, in terms of the responses of LTI systems to these idealized signals.

Chapters 3 through 6 present a thorough and self-contained development of the methods of Fourier analysis in both continuous and discrete time and together represent the most significant reorganization and revision in the second edition. In particular, as we indicated previously, we have introduced the concept of frequency-domain filtering at a much earlier point in the development in order to provide motivation for and a concrete application of the Fourier methods being developed. As in the first edition, we begin the discussions in Chapter 3 by emphasizing and illustrating the two fundamental reasons for the important role Fourier analysis plays in the study of signals and systems in both continuous and discrete time: (1) extremely broad classes of signals can be represented as weighted sums or integrals of complex exponentials; and (2) the response of an LTI system to a complex exponential input is the same exponential multiplied by a complex-number characteristic of the system. However, in contrast to the first edition, the focus of attention in Chapter 3 is on Fourier series representations for periodic signals in both continuous time and discrete time. In this way we not only introduce and examine many of the properties of Fourier representations without the additional mathematical generalization required to obtain the Fourier transform for aperiodic signals, but we also can introduce the application to filtering at a very early stage in the development. In particular, taking advantage of the fact that complex exponentials are eigenfunctions of LTI systems, we introduce the frequency response of an LTI system and use it to discuss the concept of frequency-selective filtering, to introduce ideal filters, and to give several examples of nonideal filters described by differential and difference equations. In this way, with a minimum of mathematical preliminaries, we provide the student with a deeper appreciation for what a Fourier representation means and why it is such a useful construct.

Chapters 4 and 5 then build on the foundation provided by Chapter 3 as we develop first the continuous-time Fourier transform in Chapter 4 and, in a parallel fashion, the discrete-time Fourier transform in Chapter 5. In both chapters we derive the Fourier transform representation of an aperiodic signal as the limit of the Fourier series for a signal whose period becomes arbitrarily large. This perspective emphasizes the close relationship between Fourier series and transforms, which we develop further in subsequent sections and which allows us to transfer the intuition developed for Fourier series in Chapter 3 to the more general context of Fourier transforms. In both chapters we have included a discussion of the many important properties of Fourier transforms, with special emphasis placed on the convolution and multiplication properties. In particular, the convolution property allows us to take a second look at the topic of frequency-selective filtering, while the multiplication property serves as the starting point for our treatment of sampling and modulation in later chapters. Finally, in the last sections in Chapters 4 and 5 we use transform methods to determine the frequency responses of LTI systems described by differential and difference equations and to provide several examples illustrating how Fourier transforms can be used to compute the responses for such systems. To supplement these discussions (and later treatments of Laplace and z-transforms) we have again included an Appendix at the end of the book that contains a description of the method of partial fraction expansion.

Our treatment of Fourier analysis in these two chapters is characteristic of the parallel treatment we have developed. Specifically, in our discussion in Chapter 5, we are able to build on much of the insight developed in Chapter 4 for the continuous-time case, and toward the end of Chapter 5 we emphasize the complete duality in continuous-time and discrete-time Fourier representations. In addition, we bring the special nature of each domain into sharper focus by contrasting the differences between continuous- and discrete-time Fourier analysis.

As those familiar with the first edition will note, the lengths and scopes of Chapters 4 and 5 in the second edition are considerably smaller than their first edition counterparts. This is due not only to the fact that Fourier series are now dealt with in a separate chapter but also to our moving several topics into Chapter 6. The result, we believe, has several significant benefits. First, the presentation in three shorter chapters of the basic concepts and results of Fourier analysis, together with the introduction of the concept of frequency-selective filtering, should help the student in organizing his or her understanding of this material and in developing some intuition about the frequency domain and appreciation for its potential applications. Then, with Chapters 3-5 as a foundation, we can engage in a more detailed look at a number of important topics and applications. In Chapter 6 we take a deeper look at both the time- and frequency-domain characteristics of LTI systems. For example, we introduce magnitude-phase and Bode plot representations for frequency responses and discuss the effect of frequency response phase on the time domain characteristics of the output of an LTI system. In addition, we examine the time- and frequency-domain behavior of ideal and nonideal filters and the tradeoffs between these that must be addressed in practice. We also take a careful look at first- and second-order systems and their roles as basic building blocks for more complex system synthesis and analysis in both continuous and discrete time. Finally, we discuss several other more complex examples of filters in both continuous and discrete time. These examples together with the numerous other aspects of filtering explored in the problems at the end of the chapter provide the student with some appreciation for the richness and flavor of this important subject. While each of the topics in Chapter 6 was present in the first edition, we believe that by reorganizing and collecting them in a separate chapter following the basic development of Fourier analysis, we have both simplified the introduction of this important topic in Chapters 3-5 and presented in Chapter 6 a considerably more cohesive picture of time- and frequency-domain issues.

In response to suggestions and preferences expressed by many users of the first edition we have modified notation in the discussion of Fourier transforms to be more consistent with notation most typically used for continuous-time and discrete-time Fourier transforms. Specifically, beginning with Chapter 3 we now denote the continuous-time Fourier transform as $X(j\omega)$ and the discrete-time Fourier transform as $X(e^{j\omega})$. As with all options with notation, there is not a unique best choice for the notation for Fourier transforms. However, it is our feeling, and that of many of our colleagues, that the notation used in this edition represents the preferable choice.

Our treatment of sampling in Chapter 7 is concerned primarily with the sampling theorem and its implications. However, to place this subject in perspective we begin by discussing the general concepts of representing a continuous-time signal in terms of its samples and the reconstruction of signals using interpolation. After using frequency-domain methods to derive the sampling theorem, we consider both the frequency and time domains to provide intuition concerning the phenomenon of aliasing resulting from undersampling. One of the very important uses of sampling is in the discrete-time processing of continuous-time signals, a topic that we explore at some length in this chapter. Following this, we turn to the sampling of discrete-time signals. The basic result underlying discrete-time sampling is developed in a manner that parallels that used in continuous time, and the applications of this result to problems of decimation and interpolation are described. Again a variety of other applications, in both continuous and discrete time, are addressed in the problems.

Once again the reader acquainted with our first edition will note a change, in this case involving the reversal in the order of the presentation of sampling and communications. We have chosen to place sampling before communications in the second edition both because we can call on simple intuition to motivate and describe the processes of sampling and reconstruction from samples and also because this order of presentation then allows us in Chapter 8 to talk more easily about forms of communication systems that are closely related to sampling or rely fundamentally on using a sampled version of the signal to be transmitted.

Our treatment of communications in Chapter 8 includes an in-depth discussion of continuous-time sinusoidal amplitude modulation (AM), which begins with the straightforward application of the multiplication property to describe the effect of sinusoidal AM in the frequency domain and to suggest how the original modulating signal can be recovered. Following this, we develop a number of additional issues and applications related to sinusoidal modulation, including frequency-division multiplexing and single-sideband modulation. Many other examples and applications are described in the problems. Several additional topics are covered in Chapter 8. The first of these is amplitude modulation of a pulse train and time-division multiplexing, which has a close connection to the topic of sampling in Chapter 7. Indeed we make this tie even more explicit and provide a look into the important field of digital communications by introducing and briefly describing the topics of pulse-amplitude modulation (PAM) and intersymbol interference. Finally, our discussion of frequency modulation (FM) provides the reader with a look at a nonlinear modulation problem. Although the analysis of FM systems is not as straightforward as for the AM case, our introductory treatment indicates how frequency-domain methods can be used to gain a significant amount of insight into the characteristics of FM signals and systems. Through these discussions and the many other aspects of modulation and communications explored in the problems in this chapter we believe that the student can gain an appreciation both for the richness of the field of communications and for the central role that the tools of signals and systems analysis play in it.

Chapters 9 and 10 treat the Laplace and z-transforms, respectively. For the most part, we focus on the bilateral versions of these transforms, although in the last section of each chapter we discuss unilateral transforms and their use in solving differential and difference equations with nonzero initial conditions. Both chapters include discussions on: the close relationship between these transforms and Fourier transforms; the class of rational transforms and their representation in terms of poles and zeros; the region of convergence of a Laplace or z-transform and its relationship to properties of the signal with which it is associated; inverse transforms using partial fraction expansion; the geometric evaluation of system functions and frequency responses from pole-zero plots; and basic transform properties. In addition, in each chapter we examine the properties and uses of system functions for LTI systems. Included in these discussions are the determination of system functions for systems characterized by differential and difference equations; the use of system function algebra for interconnections of LTI systems; and the construction of cascade, parallel- and direct-form block-diagram representations for systems with rational system functions.

The tools of Laplace and z-transforms form the basis for our examination of linear feedback systems in Chapter 11. We begin this chapter by describing a number of the important uses and properties of feedback systems, including stabilizing unstable systems, designing tracking systems, and reducing system sensitivity. In subsequent sections we use the tools that we have developed in previous chapters to examine three topics that are of importance for both continuous-time and discrete-time feedback systems. These are root locus analysis, Nyquist plots and the Nyquist criterion, and log-magnitude/phase plots and the concepts of phase and gain margins for stable feedback systems.

The subject of signals and systems is an extraordinarily rich one, and a variety of approaches can be taken in designing an introductory course. It was our intention with the first edition and again with this second edition to provide instructors with a great deal of flexibility in structuring their presentations of the subject. To obtain this flexibility and to maximize the usefulness of this book for instructors, we have chosen to present thorough, in-depth treatments of a cohesive set of topics that forms the core of most introductory courses on signals and systems. In achieving this depth we have of necessity omitted introductions to topics such as descriptions of random signals and state space models that are sometimes included in first courses on signals and systems. Traditionally, at many schools, such topics are not included in introductory courses but rather are developed in more depth in follow-on undergraduate courses or in courses explicitly devoted to their investigation. Although we have not included an introduction to state space in the book, instructors of introductory courses can easily incorporate it into the treatments of differential and difference equations that can be found throughout the book. In particular, the discussions in Chapters 9 and 10 on block diagram representations for systems with rational system functions and on unilateral transforms and their use in solving differential and difference equations with initial conditions form natural points of departure for the discussions of state-space representations.

A typical one-semester course at the sophomore-junior level using this book would cover Chapters 1-5 in reasonable depth (although various topics in each chapter are easily omitted at the discretion of the instructor) with selected topics chosen from the remaining chapters. For example, one possibility is to present several of the basic topics in Chapters 6-8 together with a treatment of Laplace and z-transforms and perhaps a brief introduction to the use of system function concepts to analyze feedback systems. A variety of alternate formats are possible, including one that incorporates an introduction to state space or one in which more focus is placed on continuous-time systems by de-emphasizing Chapters 5 and 10 and the discrete-time topics in Chapters 3, 7, 8, and 11.

In addition to these course formats this book can be used as the basic text for a thorough, two-semester sequence on linear systems. Alternatively, the portions of the book not used in a first course on signals and systems can, together with other sources, form the basis for a subsequent course. For example, much of the material in this book forms a direct bridge to subjects such as state space analysis, control systems, digital signal processing, communications and statistical signal processing. Consequently, a follow-on course can be constructed that uses some of the topics in this book together with supplementary material in order to provide an introduction to one or more of these advanced subjects. In fact, a new course following this model has been developed at MIT and has proven not only to be a popular course among our students but also a crucial component of our signals and systems curriculum.

As it was with the first edition, in the process of writing this book we have been fortunate to have received assistance, suggestions, and support from numerous colleagues, students and friends. The ideas and perspectives that form the heart of this book have continued to evolve as a result of our own experiences in teaching signals and systems and the influences of the many colleagues and students with whom we have worked. We would like to thank Professor Ian T. Young for his contributions to the first edition of this book and to thank and welcome Professor Hamid Nawab for the significant role he played in the development and complete restructuring of the examples and problems for this second edition. We also express our appreciation to John Buck, Michael Daniel and Andrew Singer for writing the MATLAB companion to the text. In addition, we would like to thank Jason Oppenheim for the use of one of his original photographs and Vivian Berman for her ideas and help in arriving at a cover design. Also, as indicated on the acknowledgment page that follows, we are deeply grateful to the many students and colleagues who devoted a significant number of hours to a variety of aspects of the preparation of this second edition.

We would also like to express our sincere thanks to Mr. Ray Stata and Analog Devices, Inc. for their generous and continued support of signal processing and this text through funding of the Distinguished Professor Chair in Electrical Engineering. We also thank M.I.T. for providing support and an invigorating environment in which to develop our ideas.

The encouragement, patience, technical support, and enthusiasm provided by Prentice-Hall, and in particular by Marcia Horton, Tom Robbins, Don Fowley, and their predecessors and by Ralph Pescatore of TKM Productions and the production staff at Prentice-Hall, have been crucial in making this second edition a reality.

Alan V. Oppenheim
Alan S. Willsky

Cambridge, Massachusetts

FOREWORD

The concepts of signals and systems arise in a wide variety of fields, and the ideas and techniques associated with these concepts play an important role in such diverse areas of science and technology as communications, aeronautics and astronautics, circuit design, acoustics, seismology, biomedical engineering, energy generation and distribution systems, chemical process control, and speech processing. Although the physical nature of the signals and systems that arise in these various disciplines may be drastically different, they all have two very basic features in common. The signals, which are functions of one or more independent variables, contain information about the behavior or nature of some phenomenon, whereas the systems respond to particular signals by producing other signals or some desired behavior. Voltages and currents as a function of time in an electrical circuit are examples of signals, and a circuit is itself an example of a system, which in this case responds to applied voltages and currents. As another example, when an automobile driver depresses the accelerator pedal, the automobile responds by increasing the speed of the vehicle. In this case, the system is the automobile, the pressure on the accelerator pedal the input to the system, and the automobile speed the response. A computer program for the automated diagnosis of electrocardiograms can be viewed as a system which has as its input a digitized electrocardiogram and which produces estimates of parameters such as heart rate as outputs. A camera is a system that receives light from different sources and reflected from objects and produces a photograph. A robot arm is a system whose movements are the response to control inputs.

In the many contexts in which signals and systems arise, there are a variety of problems and questions that are of importance. In some cases, we are presented with a specific system and are interested in characterizing it in detail to understand how it will respond to various inputs. Examples include the analysis of a circuit in order to quantify its response to different voltage and current sources and the determination of an aircraft's response characteristics both to pilot commands and to wind gusts.

In other problems of signal and system analysis, rather than analyzing existing systems, our interest may be focused on designing systems to process signals in particular ways. One very common context in which such problems arise is in the design of systems to enhance or restore signals that have been degraded in some way. For example, when a pilot is communicating with an air traffic control tower, the communication can be degraded by the high level of background noise in the cockpit. In this and many similar cases, it is possible to design systems that will retain the desired signal, in this case the pilot's voice, and reject (at least approximately) the unwanted signal, i.e., the noise. A similar set of objectives can also be found in the general area of image restoration and image enhancement. For example, images from deep space probes or earth-observing satellites typically represent degraded versions of the scenes being imaged because of limitations of the imaging equipment, atmospheric effects, and errors in signal transmission in returning the images to earth. Consequently, images returned from space are routinely processed by systems to compensate for some of these degradations. In addition, such images are usually processed to enhance certain features, such as lines (corresponding, for example, to river beds or faults) or regional boundaries in which there are sharp contrasts in color or darkness.

In addition to enhancement and restoration, in many applications there is a need to design systems to extract specific pieces of information from signals. The estimation of heart rate from an electrocardiogram is one example. Another arises in economic forecasting. We may, for example, wish to analyze the history of an economic time series, such as a set of stock market averages, in order to estimate trends and other characteristics such as seasonal variations that may be of use in making predictions about future behavior. In other applications, the focus may be on the design of signals with particular properties. Specifically, in communications applications considerable attention is paid to designing signals to meet the constraints and requirements for successful transmission. For example, long distance communication through the atmosphere requires the use of signals with frequencies in a particular part of the electromagnetic spectrum. The design of communication signals must also take into account the need for reliable reception in the presence of both distortion due to transmission through the atmosphere and interference from other signals being transmitted simultaneously by other users.

Another very important class of applications in which the concepts and techniques of signal and system analysis arise are those in which we wish to modify or control the characteristics of a given system, perhaps through the choice of specific input signals or by combining the system with other systems. Illustrative of this kind of application is the design of control systems to regulate chemical processing plants. Plants of this type are equipped with a variety of sensors that measure physical signals such as temperature, humidity, and chemical composition. The control system in such a plant responds to these sensor signals by adjusting quantities such as flow rates and temperature in order to regulate the ongoing chemical process. The design of aircraft autopilots and computer control systems represents another example. In this case, signals measuring aircraft speed, altitude, and heading are used by the aircraft's control system in order to adjust variables such as throttle setting and the position of the rudder and ailerons. These adjustments are made to ensure that the aircraft follows a specified course, to smooth out the aircraft's ride, and to enhance its responsiveness to pilot commands. In both this case and in the previous example of chemical process control, an important concept, referred to as feedback, plays a major role, as measured signals are fed back and used to adjust the response characteristics of a system.

The examples in the preceding paragraphs represent only a few of an extraordinarily wide variety of applications for the concepts of signals and systems. The importance of these concepts stems not only from the diversity of phenomena and processes in which they arise, but also from the collection of ideas, analytical techniques, and methodologies that have been and are being developed and used to solve problems involving signals and systems. The history of this development extends back over many centuries, and although most of this work was motivated by specific applications, many of these ideas have proven to be of central importance to problems in a far larger variety of contexts than those for which they were originally intended. For example, the tools of Fourier analysis, which form the basis for the frequency-domain analysis of signals and systems, and which we will develop in some detail in this book, can be traced from problems of astronomy studied by the ancient Babylonians to the development of mathematical physics in the eighteenth and nineteenth centuries.

In some of the examples that we have mentioned, the signals vary continuously in time, whereas in others, their evolution is described only at discrete points in time. For example, in the analysis of electrical circuits and mechanical systems we are concerned with signals that vary continuously. On the other hand, the daily closing stock market average is by its very nature a signal that evolves at discrete points in time (i.e., at the close of each day). Rather than a curve as a function of a continuous variable, then, the closing stock market average is a sequence of numbers associated with the discrete time instants at which it is specified. This distinction in the basic description of the evolution of signals and of the systems that respond to or process these signals leads naturally to two parallel frameworks for signal and system analysis, one for phenomena and processes that are described in continuous time and one for those that are described in discrete time.

The concepts and techniques associated both with continuous-time signals and systems and with discrete-time signals and systems have a rich history and are conceptually closely related. Historically, however, because their applications have in the past been sufficiently different, they have for the most part been studied and developed somewhat separately. Continuous-time signals and systems have very strong roots in problems associated with physics and, in the more recent past, with electrical circuits and communications. The techniques of discrete-time signals and systems have strong roots in numerical analysis, statistics, and time-series analysis associated with such applications as the analysis of economic and demographic data. Over the past several decades, however, the disciplines of continuous-time and discrete-time signals and systems have become increasingly entwined and the applications have become highly interrelated. The major impetus for this has come from the dramatic advances in technology for the implementation of systems and for the generation of signals. Specifically, the continuing development of high-speed digital computers, integrated circuits, and sophisticated high-density device fabrication techniques has made it increasingly advantageous to consider processing continuous-time signals by representing them by time samples (i.e., by converting them to discrete-time signals). As one example, the computer control system for a modern high-performance aircraft digitizes sensor outputs such as vehicle speed in order to produce a sequence of sampled measurements which are then processed by the control system.

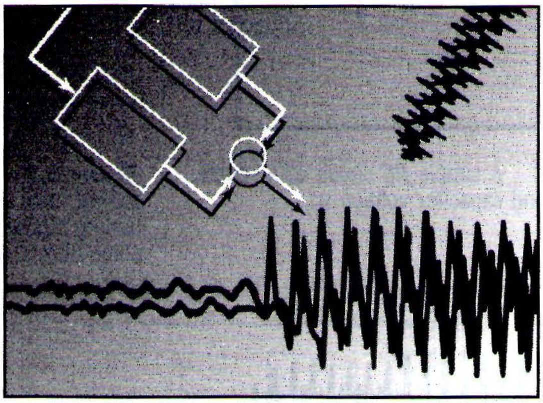
Because of the growing interrelationship between continuous-time signals and systems and discrete-time signals and systems and because of the close relationship among the concepts and techniques associated with each, we have chosen in this text to develop the concepts of continuous-time and discrete-time signals and systems in parallel. Since many of the concepts are similar (but not identical), by treating them in parallel, insight and intuition can be shared and both the similarities and differences between them become better focused. In addition, as will be evident as we proceed through the material, there are some concepts that are inherently easier to understand in one framework than the other and, once understood, the insight is easily transferable. Furthermore, this parallel treatment greatly facilitates our understanding of the very important practical context in which continuous and discrete time are brought together, namely the sampling of continuous-time signals and the processing of continuous-time signals using discrete-time systems.

As we have so far described them, the notions of signals and systems are extremely general concepts. At this level of generality, however, only the most sweeping statements can be made about the nature of signals and systems, and their properties can be discussed only in the most elementary terms. On the other hand, an important and fundamental notion in dealing with signals and systems is that by carefully choosing subclasses of each with particular properties that can then be exploited, we can analyze and characterize these signals and systems in great depth. The principal focus in this book is on the particular class of linear time-invariant systems. The properties of linearity and time invariance that define this class lead to a remarkable set of concepts and techniques which are not only of major practical importance but also analytically tractable and intellectually satisfying.

As we have emphasized in this foreword, signal and system analysis has a long history out of which have emerged some basic techniques and fundamental principles which have extremely broad areas of application. Indeed, signal and system analysis is constantly evolving and developing in response to new problems, techniques, and opportunities. We fully expect this development to accelerate in pace as improved technology makes possible the implementation of increasingly complex systems and signal processing techniques. In the future we will see signals and systems tools and concepts applied to an expanding scope of applications. For these reasons, we feel that the topic of signal and system analysis represents a body of knowledge that is of essential concern to the scientist and engineer. We have chosen the set of topics presented in this book, the organization of the presentation, and the problems in each chapter in a way that we feel will most help the reader to obtain a solid foundation in the fundamentals of signal and system analysis; to gain an understanding of some of the very important and basic applications of these fundamentals to problems in filtering, sampling, communications, and feedback system analysis; and to develop some appreciation for an extremely powerful and broadly applicable approach to formulating and solving complex problems.

1

SIGNALS AND SYSTEMS



1.0 INTRODUCTION

As described in the Foreword, the intuitive notions of signals and systems arise in a rich variety of contexts. Moreover, as we will see in this book, there is an analytical framework—that is, a language for describing signals and systems and an extremely powerful set of tools for analyzing them—that applies equally well to problems in many fields. In this chapter, we begin our development of the analytical framework for signals and systems by introducing their mathematical description and representations. In the chapters that follow, we build on this foundation in order to develop and describe additional concepts and methods that add considerably both to our understanding of signals and systems and to our ability to analyze and solve problems involving signals and systems that arise in a broad array of applications.

1.1 CONTINUOUS-TIME AND DISCRETE-TIME SIGNALS

1.1.1 Examples and Mathematical Representation

Signals may describe a wide variety of physical phenomena. Although signals can be represented in many ways, in all cases the information in a signal is contained in a pattern of variations of some form. For example, consider the simple circuit in Figure 1.1. In this case, the patterns of variation over time in the source and capacitor voltages, v_s and v_c, are examples of signals. Similarly, as depicted in Figure 1.2, the variations over time of the applied force f and the resulting automobile velocity v are signals. As another example, consider the human vocal mechanism, which produces speech by creating fluctuations in acoustic pressure. Figure 1.3 is an illustration of a recording of such a speech signal, obtained by

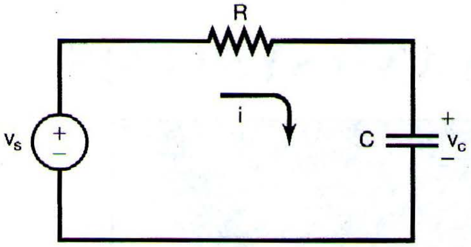


Figure 1.1 A simple *RC* circuit with source voltage v_s and capacitor voltage v_c.

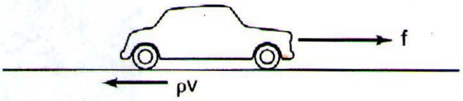


Figure 1.2 An automobile responding to an applied force *f* from the engine and to a retarding frictional force ρv proportional to the automobile's velocity *v*.

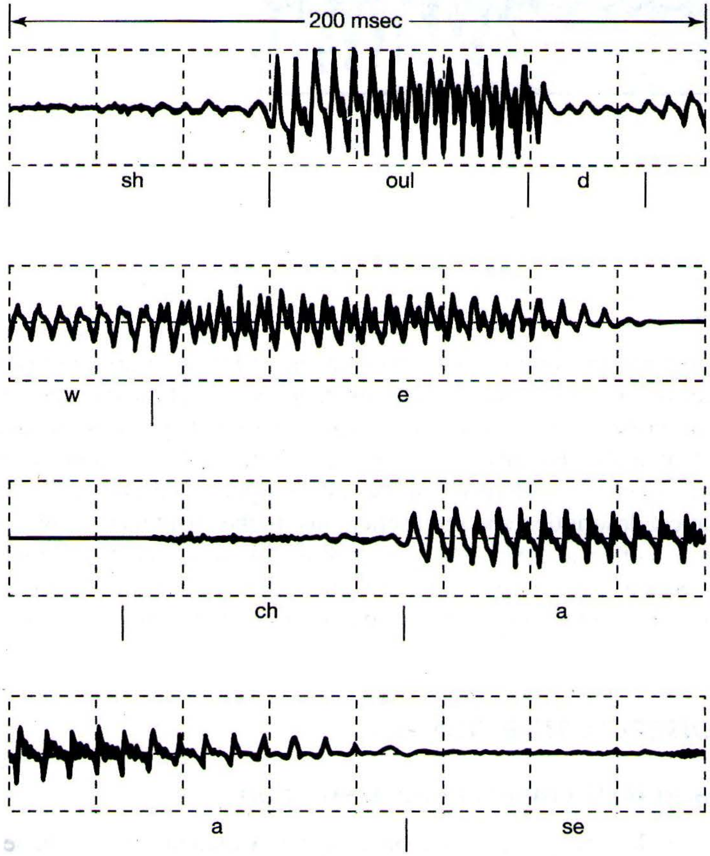


Figure 1.3 Example of a recording of speech. [Adapted from *Applications of Digital Signal Processing*, A.V. Oppenheim, ed. (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1978), p. 121.] The signal represents acoustic pressure variations as a function of time for the spoken words "should we chase." The top line of the figure corresponds to the word "should," the second line to the word "we," and the last two lines to the word "chase." (We have indicated the approximate beginnings and endings of each successive sound in each word.)

using a microphone to sense variations in acoustic pressure, which are then converted into an electrical signal. As can be seen in the figure, different sounds correspond to different patterns in the variations of acoustic pressure, and the human vocal system produces intelligible speech by generating particular sequences of these patterns. Alternatively, for the monochromatic picture, shown in Figure 1.4, it is the pattern of variations in brightness across the image that is important.

Figure 1.4 A monochromatic picture.

Signals are represented mathematically as functions of one or more independent variables. For example, a speech signal can be represented mathematically by acoustic pressure as a function of time, and a picture can be represented by brightness as a function of two spatial variables. In this book, we focus our attention on signals involving a single independent variable. For convenience, we will generally refer to the independent variable as time, although it may not in fact represent time in specific applications. For example, in geophysics, signals representing variations with depth of physical quantities such as density, porosity, and electrical resistivity are used to study the structure of the earth. Also, knowledge of the variations of air pressure, temperature, and wind speed with altitude are extremely important in meteorological investigations. Figure 1.5 depicts a typical example of annual average vertical wind profile as a function of height. The measured variations of wind speed with height are used in examining weather patterns, as well as wind conditions that may affect an aircraft during final approach and landing.

Throughout this book we will be considering two basic types of signals: continuous-time signals and discrete-time signals. In the case of continuous-time signals the independent variable is continuous, and thus these signals are defined for a continuum of values

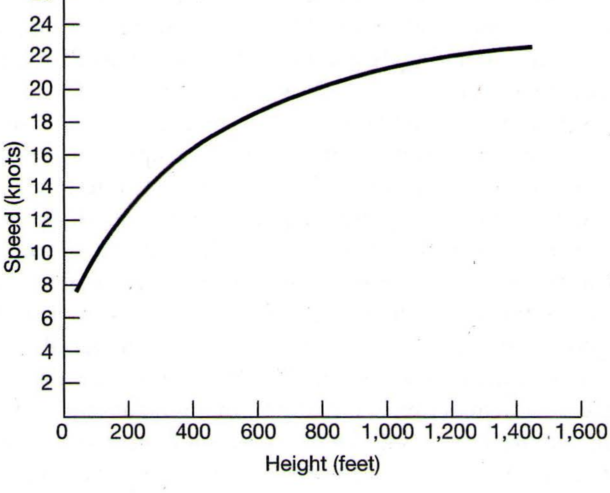


Figure 1.5 Typical annual vertical wind profile. (Adapted from Crawford and Hudson, National Severe Storms Laboratory Report, ESSA ERLTM-NSSL 48, August 1970.)

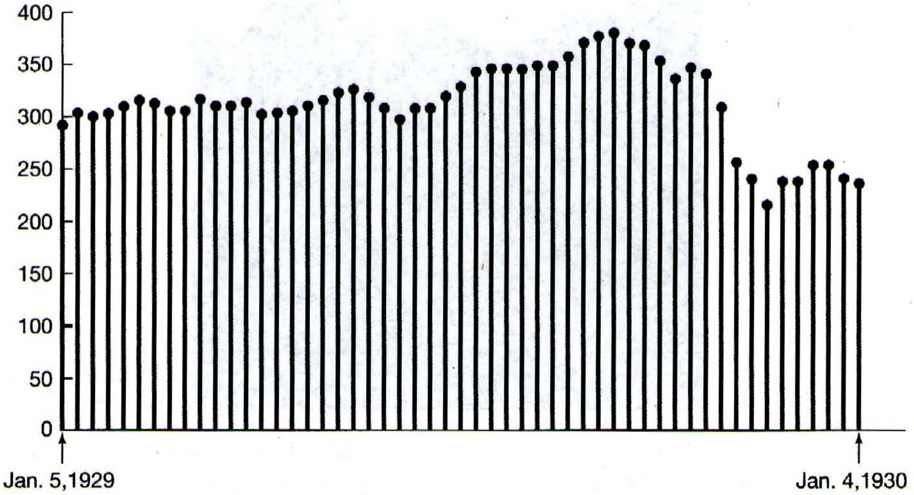


Figure 1.6 An example of a discrete-time signal: The weekly Dow-Jones stock market index from January 5, 1929, to January 4, 1930.

of the independent variable. On the other hand, discrete-time signals are defined only at discrete times, and consequently, for these signals, the independent variable takes on only a discrete set of values. A speech signal as a function of time and atmospheric pressure as a function of altitude are examples of continuous-time signals. The weekly Dow-Jones stock market index, as illustrated in Figure 1.6, is an example of a discrete-time signal. Other examples of discrete-time signals can be found in demographic studies in which various attributes, such as average budget, crime rate, or pounds of fish caught, are tabulated against such discrete variables as family size, total population, or type of fishing vessel, respectively.

To distinguish between continuous-time and discrete-time signals, we will use the symbol t to denote the continuous-time independent variable and n to denote the discrete-time independent variable. In addition, for continuous-time signals we will enclose the independent variable in parentheses $(\cdot)$, whereas for discrete-time signals we will use brackets $[\cdot]$ to enclose the independent variable. We will also have frequent occasions when it will be useful to represent signals graphically. Illustrations of a continuous-time signal $x(t)$ and a discrete-time signal $x[n]$ are shown in Figure 1.7. It is important to note that the discrete-time signal $x[n]$ is defined *only* for integer values of the independent variable. Our choice of graphical representation for $x[n]$ emphasizes this fact, and for further emphasis we will on occasion refer to $x[n]$ as a discrete-time *sequence*.

A discrete-time signal $x[n]$ may represent a phenomenon for which the independent variable is inherently discrete. Signals such as demographic data are examples of this. On the other hand, a very important class of discrete-time signals arises from the *sampling* of continuous-time signals. In this case, the discrete-time signal $x[n]$ represents successive samples of an underlying phenomenon for which the independent variable is continuous. Because of their speed, computational power, and flexibility, modern digital processors are used to implement many practical systems, ranging from digital autopilots to digital audio systems. Such systems require the use of discrete-time sequences representing sampled versions of continuous-time signals—e.g., aircraft position, velocity, and heading for an

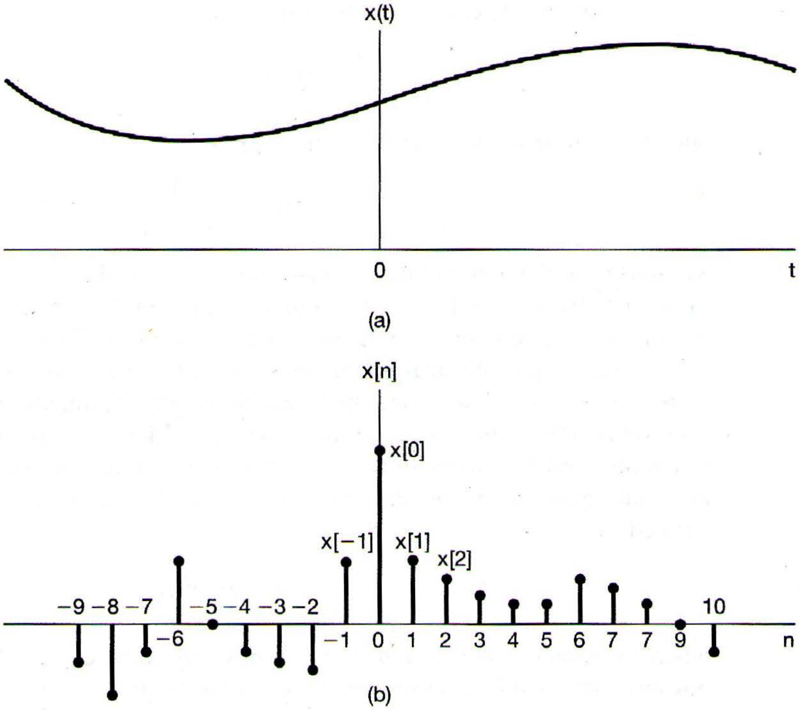


Figure 1.7 Graphical representations of (a) continuous-time and (b) discrete-time signals.

autopilot or speech and music for an audio system. Also, pictures in newspapers—or in this book, for that matter—actually consist of a very fine grid of points, and each of these points represents a sample of the brightness of the corresponding point in the original image. No matter what the source of the data, however, the signal $x[n]$ is defined only for integer values of n. It makes no more sense to refer to the $3\frac{1}{2}$th sample of a digital speech signal than it does to refer to the average budget for a family with $2\frac{1}{2}$ family members.

Throughout most of this book we will treat discrete-time signals and continuous-time signals separately but in parallel, so that we can draw on insights developed in one setting to aid our understanding of another. In Chapter 7 we will return to the question of sampling, and in that context we will bring continuous-time and discrete-time concepts together in order to examine the relationship between a continuous-time signal and a discrete-time signal obtained from it by sampling.

1.1.2 Signal Energy and Power

From the range of examples provided so far, we see that signals may represent a broad variety of phenomena. In many, but not all, applications, the signals we consider are directly related to physical quantities capturing power and energy in a physical system. For example, if $v(t)$ and $i(t)$ are, respectively, the voltage and current across a resistor with resistance R, then the instantaneous power is

$$p(t) = v(t)i(t) = \frac{1}{R}v^2(t). \tag{1.1}$$

The total *energy* expended over the time interval $t_1 \le t \le t_2$ is

$$\int_{t_1}^{t_2} p(t)\,dt = \int_{t_1}^{t_2} \frac{1}{R}v^2(t)\,dt, \tag{1.2}$$

and the *average power* over this time interval is

$$\frac{1}{t_2 - t_1}\int_{t_1}^{t_2} p(t)\,dt = \frac{1}{t_2 - t_1}\int_{t_1}^{t_2} \frac{1}{R}v^2(t)\,dt. \tag{1.3}$$

Similarly, for the automobile depicted in Figure 1.2, the instantaneous power dissipated through friction is $p(t) = bv^2(t)$, and we can then define the total energy and average power over a time interval in the same way as in eqs. (1.2) and (1.3).

With simple physical examples such as these as motivation, it is a common and worthwhile convention to use similar terminology for power and energy for *any* continuous-time signal $x(t)$ or *any* discrete-time signal $x[n]$. Moreover, as we will see shortly, we will frequently find it convenient to consider signals that take on complex values. In this case, the total energy over the time interval $t_1 \le t \le t_2$ in a continuous-time signal $x(t)$ is defined as

$$\int_{t_1}^{t_2} |x(t)|^2 dt, \tag{1.4}$$

where $|x|$ denotes the magnitude of the (possibly complex) number x. The time-averaged power is obtained by dividing eq. (1.4) by the length, $t_2 - t_1$, of the time interval. Similarly, the total energy in a discrete-time signal $x[n]$ over the time interval $n_1 \le n \le n_2$ is defined as

$$\sum_{n=n_1}^{n_2} |x[n]|^2, \tag{1.5}$$

and dividing by the number of points in the interval, $n_2 - n_1 + 1$, yields the average power over the interval. It is important to remember that the terms "power" and "energy" are used here independently of whether the quantities in eqs. (1.4) and (1.5) actually are related to physical energy.[1] Nevertheless, we will find it convenient to use these terms in a general fashion.

Furthermore, in many systems we will be interested in examining power and energy in signals over an infinite time interval, i.e., for $-\infty < t < +\infty$ or for $-\infty < n < +\infty$. In these cases, we define the total energy as limits of eqs. (1.4) and (1.5) as the time interval increases without bound. That is, in continuous time,

$$E_\infty \triangleq \lim_{T\to\infty}\int_{-T}^{T} |x(t)|^2\,dt = \int_{-\infty}^{+\infty} |x(t)|^2\,dt, \tag{1.6}$$

and in discrete time,

$$E_\infty \triangleq \lim_{N\to\infty}\sum_{n=-N}^{+N} |x[n]|^2 = \sum_{n=-\infty}^{+\infty} |x[n]|^2. \tag{1.7}$$

---

[1] Even if such a relationship does exist, eqs. (1.4) and (1.5) may have the wrong dimensions and scalings. For example, comparing eqs. (1.2) and (1.4), we see that if $x(t)$ represents the voltage across a resistor, then eq. (1.4) must be divided by the resistance (measured, for example, in ohms) to obtain units of physical energy.

Note that for some signals the integral in eq. (1.6) or sum in eq. (1.7) might not converge—e.g., if $x(t)$ or $x[n]$ equals a nonzero constant value for all time. Such signals have infinite energy, while signals with $E_\infty < \infty$ have finite energy.

In an analogous fashion, we can define the time-averaged power over an infinite interval as

$$P_\infty \triangleq \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} |x(t)|^2\, dt \tag{1.8}$$

and

$$P_\infty \triangleq \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{+N} |x[n]|^2 \tag{1.9}$$

in continuous time and discrete time, respectively. With these definitions, we can identify three important classes of signals. The first of these is the class of signals with finite total energy, i.e., those signals for which $E_\infty < \infty$. Such a signal must have zero average power, since in the continuous time case, for example, we see from eq. (1.8) that

$$P_\infty = \lim_{T\to\infty} \frac{E_\infty}{2T} = 0. \tag{1.10}$$

An example of a finite-energy signal is a signal that takes on the value 1 for $0 \leq t \leq 1$ and 0 otherwise. In this case, $E_\infty = 1$ and $P_\infty = 0$.

A second class of signals are those with finite average power P_∞. From what we have just seen, if $P_\infty > 0$, then, of necessity, $E_\infty = \infty$. This, of course, makes sense, since if there is a nonzero average energy per unit time (i.e., nonzero power), then integrating or summing this over an infinite time interval yields an infinite amount of energy. For example, the constant signal $x[n] = 4$ has infinite energy, but average power $P_\infty = 16$. There are also signals for which neither P_∞ nor E_∞ are finite. A simple example is the signal $x(t) = t$. We will encounter other examples of signals in each of these classes in the remainder of this and the following chapters.

1.2 TRANSFORMATIONS OF THE INDEPENDENT VARIABLE

A central concept in signal and system analysis is that of the transformation of a signal. For example, in an aircraft control system, signals corresponding to the actions of the pilot are transformed by electrical and mechanical systems into changes in aircraft thrust or the positions of aircraft control surfaces such as the rudder or ailerons, which in turn are transformed through the dynamics and kinematics of the vehicle into changes in aircraft velocity and heading. Also, in a high-fidelity audio system, an input signal representing music as recorded on a cassette or compact disc is modified in order to enhance desirable characteristics, to remove recording noise, or to balance the several components of the signal (e.g., treble and bass). In this section, we focus on a very limited but important class of elementary signal transformations that involve simple modification of the independent variable, i.e., the time axis. As we will see in this and subsequent sections of this chapter, these elementary transformations allow us to introduce several basic properties of signals and systems. In later chapters, we will find that they also play an important role in defining and characterizing far richer and important classes of systems.

1.2.1 Examples of Transformations of the Independent Variable

A simple and very important example of transforming the independent variable of a signal is a *time shift.* A time shift in discrete time is illustrated in Figure 1.8, in which we have two signals $x[n]$ and $x[n - n_0]$ that are identical in shape, but that are displaced or shifted relative to each other. We will also encounter time shifts in continuous time, as illustrated in Figure 1.9, in which $x(t - t_0)$ represents a delayed (if t_0 is positive) or advanced (if t_0 is negative) version of $x(t)$. Signals that are related in this fashion arise in applications such as radar, sonar, and seismic signal processing, in which several receivers at different locations observe a signal being transmitted through a medium (water, rock, air, etc.). In this case, the difference in propagation time from the point of origin of the transmitted signal to any two receivers results in a time shift between the signals at the two receivers.

A second basic transformation of the time axis is that of *time reversal.* For example, as illustrated in Figure 1.10, the signal $x[-n]$ is obtained from the signal $x[n]$ by a reflection about $n = 0$ (i.e., by reversing the signal). Similarly, as depicted in Figure 1.11, the signal $x(-t)$ is obtained from the signal $x(t)$ by a reflection about $t = 0$. Thus, if $x(t)$ represents an audio tape recording, then $x(-t)$ is the same tape recording played backward. Another transformation is that of *time scaling.* In Figure 1.12 we have illustrated three signals, $x(t)$, $x(2t)$, and $x(t/2)$, that are related by linear scale changes in the independent variable. If we again think of the example of $x(t)$ as a tape recording, then $x(2t)$ is that recording played at twice the speed, and $x(t/2)$ is the recording played at half-speed.

It is often of interest to determine the effect of transforming the independent variable of a given signal $x(t)$ to obtain a signal of the form $x(\alpha t + \beta)$, where α and β are given numbers. Such a transformation of the independent variable preserves the shape of $x(t)$, except that the resulting signal may be linearly stretched if $|\alpha| < 1$, linearly compressed if $|\alpha| > 1$, reversed in time if $\alpha < 0$, and shifted in time if β is nonzero. This is illustrated in the following set of examples.

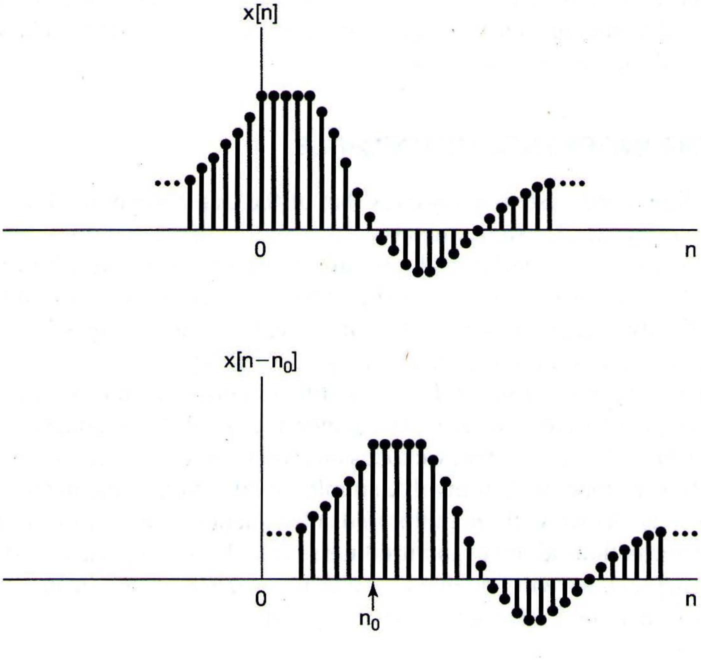


Figure 1.8 Discrete-time signals related by a time shift. In this figure $n_0 > 0$, so that $x[n - n_0]$ is a delayed verson of $x[n]$ (i.e., each point in $x[n]$ occurs later in $x[n - n_0]$).

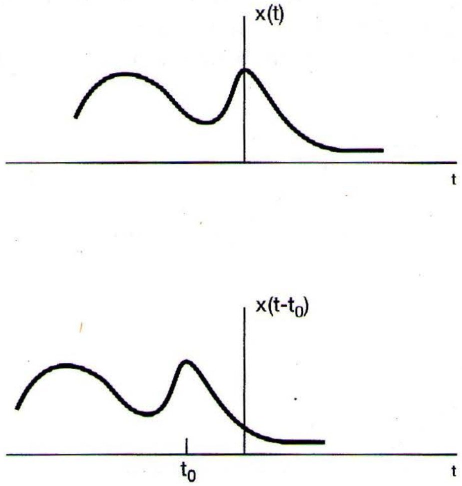

x(t)
t
x(t-t0)
t0
t

Figure 1.9 Continuous-time signals related by a time shift. In this figure $t_0 < 0$, so that $x(t - t_0)$ is an advanced version of $x(t)$ (i.e., each point in $x(t)$ occurs at an earlier time in $x(t - t_0)$).

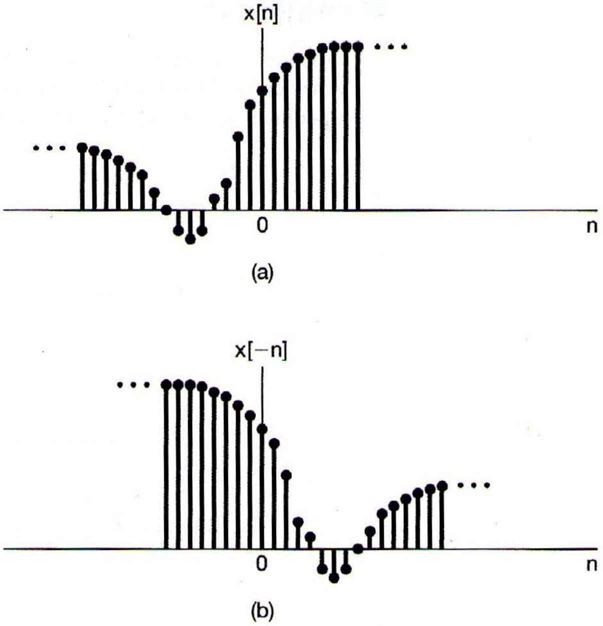

x[n]
0
n
(a)
x[−n]
0
n
(b)

Figure 1.10 (a) A discrete-time signal $x[n]$; (b) its reflection $x[-n]$ about $n = 0$.

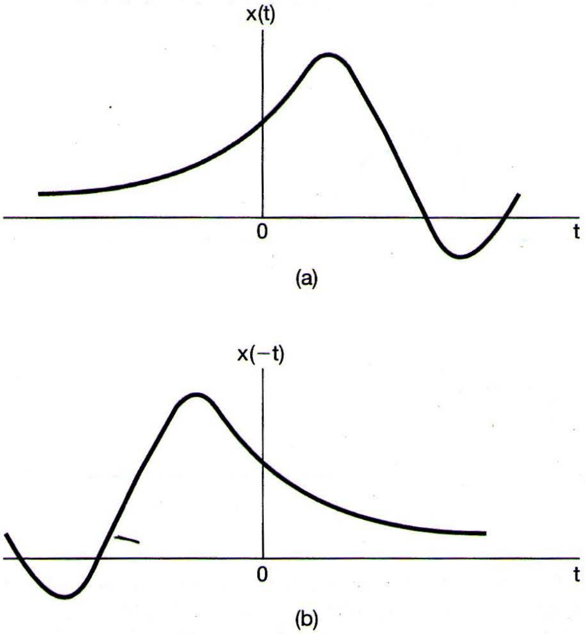

x(t)
0
t
(a)
x(−t)
0
t
(b)

Figure 1.11 (a) A continuous-time signal $x(t)$; (b) its reflection $x(-t)$ about $t = 0$.

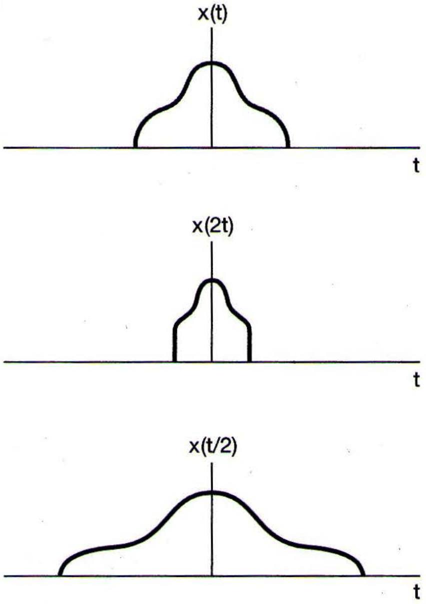

x(t)
t
x(2t)
t
x(t/2)
t

Figure 1.12 Continuous-time signals related by time scaling.

Example 1.1

Given the signal $x(t)$ shown in Figure 1.13(a), the signal $x(t+1)$ corresponds to an advance (shift to the left) by one unit along the t axis as illustrated in Figure 1.13(b). Specifically, we note that the value of $x(t)$ at $t = t_0$ occurs in $x(t+1)$ at $t = t_0 - 1$. For

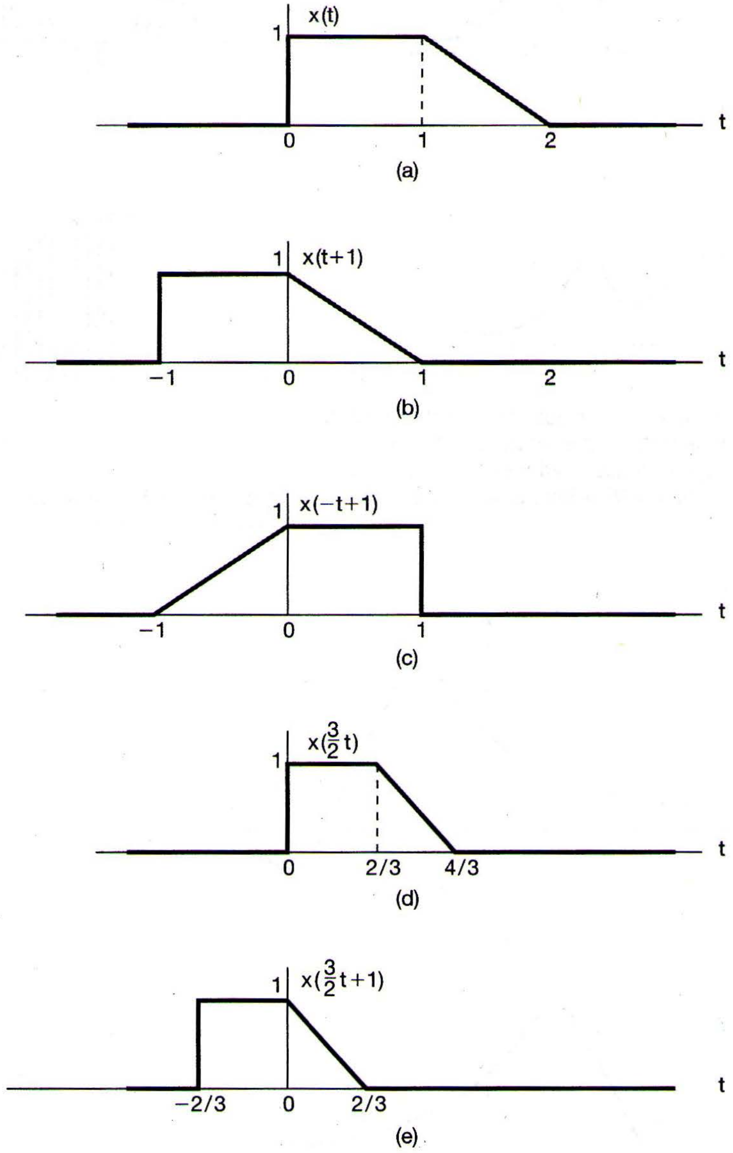


Figure 1.13 (a) The continuous-time signal $x(t)$ used in Examples 1.1–1.3 to illustrate transformations of the independent variable; (b) the time-shifted signal $x(t+1)$; (c) the signal $x(-t+1)$ obtained by a time shift and a time reversal; (d) the time-scaled signal $x(\frac{3}{2}t)$; and (e) the signal $x(\frac{3}{2}t+1)$ obtained by time-shifting and scaling.

example, the value of $x(t)$ at $t = 1$ is found in $x(t+1)$ at $t = 1-1 = 0$. Also, since $x(t)$ is zero for $t < 0$, we have $x(t+1)$ zero for $t < -1$. Similarly, since $x(t)$ is zero for $t > 2$, $x(t+1)$ is zero for $t > 1$.

Let us also consider the signal $x(-t+1)$, which may be obtained by replacing t with $-t$ in $x(t+1)$. That is, $x(-t+1)$ is the time reversed version of $x(t+1)$. Thus, $x(-t+1)$ may be obtained graphically by reflecting $x(t+1)$ about the t axis as shown in Figure 1.13(c).

Example 1.2

Given the signal $x(t)$, shown in Figure 1.13(a), the signal $x(\frac{3}{2}t)$ corresponds to a linear compression of $x(t)$ by a factor of $\frac{2}{3}$ as illustrated in Figure 1.13(d). Specifically we note that the value of $x(t)$ at $t = t_0$ occurs in $x(\frac{3}{2}t)$ at $t = \frac{2}{3}t_0$. For example, the value of $x(t)$ at $t = 1$ is found in $x(\frac{3}{2}t)$ at $t = \frac{2}{3}(1) = \frac{2}{3}$. Also, since $x(t)$ is zero for $t < 0$, we have $x(\frac{3}{2}t)$ zero for $t < 0$. Similarly, since $x(t)$ is zero for $t > 2$, $x(\frac{3}{2}t)$ is zero for $t > \frac{4}{3}$.

Example 1.3

Suppose that we would like to determine the effect of transforming the independent variable of a given signal, $x(t)$, to obtain a signal of the form $x(\alpha t + \beta)$, where α and β are given numbers. A systematic approach to doing this is to first delay or advance $x(t)$ in accordance with the value of β, and then to perform time scaling and/or time reversal on the resulting signal in accordance with the value of α. The delayed or advanced signal is linearly stretched if $|\alpha| < 1$, linearly compressed if $|\alpha| > 1$, and reversed in time if $\alpha < 0$.

To illustrate this approach, let us show how $x(\frac{3}{2}t + 1)$ may be determined for the signal $x(t)$ shown in Figure 1.13(a). Since $\beta = 1$, we first advance (shift to the left) $x(t)$ by 1 as shown in Figure 1.13(b). Since $|\alpha| = \frac{3}{2}$, we may linearly compress the shifted signal of Figure 1.13(b) by a factor of $\frac{2}{3}$ to obtain the signal shown in Figure 1.13(e).

In addition to their use in representing physical phenomena such as the time shift in a sonar signal and the speeding up or reversal of an audiotape, transformations of the independent variable are extremely useful in signal and system analysis. In Section 1.6 and in Chapter 2, we will use transformations of the independent variable to introduce and analyze the properties of systems. These transformations are also important in defining and examining some important properties of signals.

1.2.2 Periodic Signals

An important class of signals that we will encounter frequently throughout this book is the class of *periodic* signals. A periodic continuous-time signal $x(t)$ has the property that there is a positive value of T for which

$$x(t) = x(t+T) \tag{1.11}$$

for all values of t. In other words, a periodic signal has the property that it is unchanged by a time shift of T. In this case, we say that $x(t)$ is *periodic with period* T. Periodic continuous-time signals arise in a variety of contexts. For example, as illustrated in Problem 2.61, the natural response of systems in which energy is conserved, such as ideal *LC* circuits without resistive energy dissipation and ideal mechanical systems without frictional losses, are periodic and, in fact, are composed of some of the basic periodic signals that we will introduce in Section 1.3.

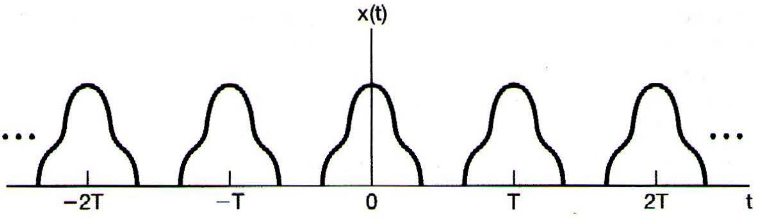


Figure 1.14 A continuous-time periodic signal.

An example of a periodic continuous-time signal is given in Figure 1.14. From the figure or from eq. (1.11), we can readily deduce that if $x(t)$ is periodic with period T, then $x(t) = x(t + mT)$ for all t and for any integer m. Thus, $x(t)$ is also periodic with period $2T, 3T, 4T, \ldots$. The *fundamental period* T_0 of $x(t)$ is the smallest positive value of T for which eq. (1.11) holds. This definition of the fundamental period works, except if $x(t)$ is a constant. In this case the fundamental period is undefined, since $x(t)$ is periodic for *any* choice of T (so there is no smallest positive value). A signal $x(t)$ that is not periodic will be referred to as an *aperiodic* signal.

Periodic signals are defined analogously in discrete time. Specifically, a discrete-time signal $x[n]$ is periodic with period N, where N is a positive integer, if it is unchanged by a time shift of N, i.e., if

$$x[n] = x[n + N] \tag{1.12}$$

for all values of n. If eq. (1.12) holds, then $x[n]$ is also periodic with period $2N, 3N, \ldots$. The *fundamental period* N_0 is the smallest positive value of N for which eq. (1.12) holds. An example of a discrete-time periodic signal with fundamental period $N_0 = 3$ is shown in Figure 1.15.

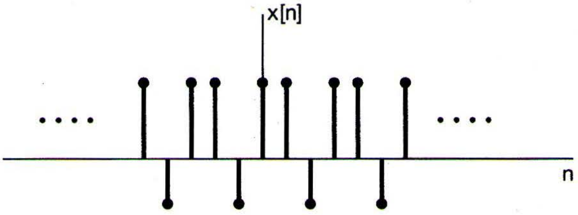


Figure 1.15 A discrete-time periodic signal with fundamental period $N_0 = 3$.

Example 1.4

Let us illustrate the type of problem solving that may be required in determining whether or not a given signal is periodic. The signal whose periodicity we wish to check is given by

$$x(t) = \begin{cases} \cos(t) & \text{if } t < 0 \\ \sin(t) & \text{if } t \geq 0 \end{cases}. \tag{1.13}$$

From trigonometry, we know that $\cos(t + 2\pi) = \cos(t)$ and $\sin(t + 2\pi) = \sin(t)$. Thus, considering $t > 0$ and $t < 0$ separately, we see that $x(t)$ does repeat itself over every interval of length 2π. However, as illustrated in Figure 1.16, $x(t)$ also has a discontinuity at the time origin that does not recur at any other time. Since every feature in the shape of a periodic signal *must* recur periodically, we conclude that the signal $x(t)$ is not periodic.

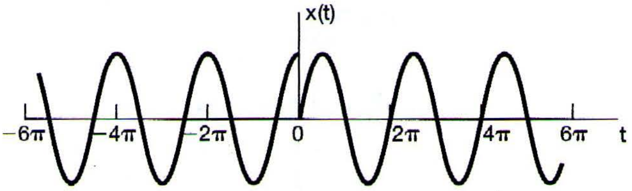


Figure 1.16 The signal $x(t)$ considered in Example 1.4.

1.2.3 Even and Odd Signals

Another set of useful properties of signals relates to their symmetry under time reversal. A signal $x(t)$ or $x[n]$ is referred to as an *even* signal if it is identical to its time-reversed counterpart, i.e., with its reflection about the origin. In continuous time a signal is even if

$$x(-t) = x(t), \tag{1.14}$$

while a discrete-time signal is even if

$$x[-n] = x[n]. \tag{1.15}$$

A signal is referred to as *odd* if

$$x(-t) = -x(t), \tag{1.16}$$

$$x[-n] = -x[n]. \tag{1.17}$$

An odd signal must necessarily be 0 at $t = 0$ or $n = 0$, since eqs. (1.16) and (1.17) require that $x(0) = -x(0)$ and $x[0] = -x[0]$. Examples of even and odd continuous-time signals are shown in Figure 1.17.

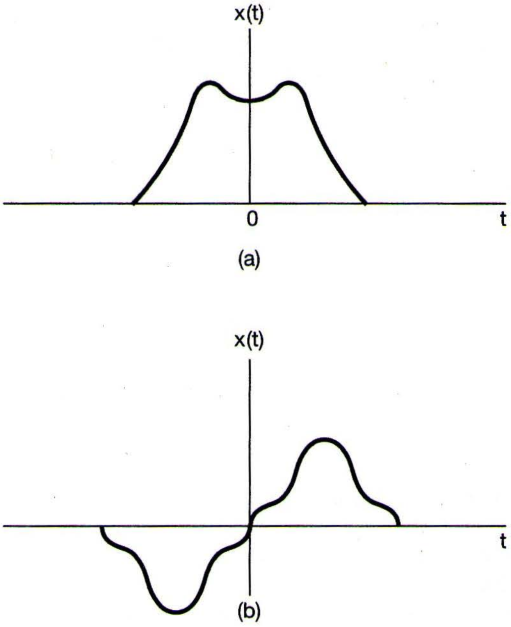


Figure 1.17 (a) An even continuous-time signal; (b) an odd continuous-time signal.

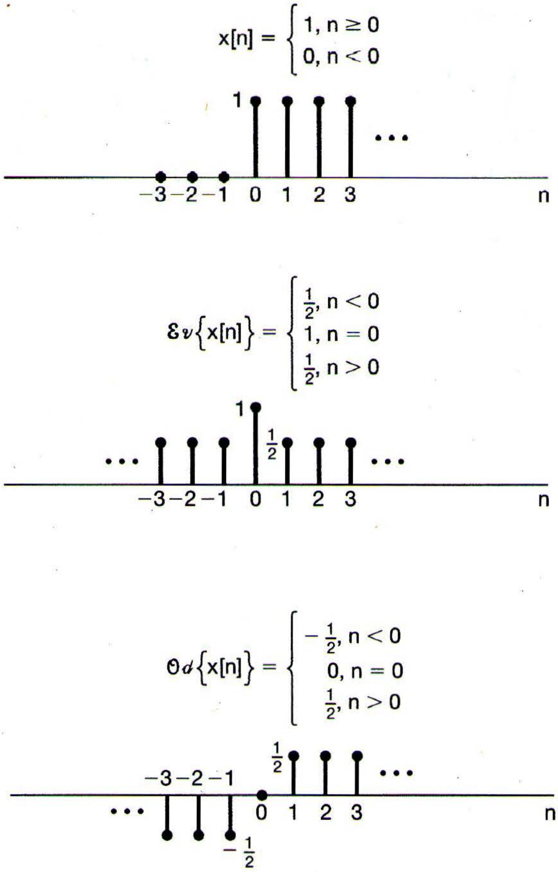

x[n] = { 1, n ≥ 0; 0, n < 0
1
...
−3 −2 −1 0 1 2 3
n
εv{x[n]} = { 1/2, n < 0; 1, n = 0; 1/2, n > 0
1
1/2
...
...
−3 −2 −1 0 1 2 3
n
Od{x[n]} = { −1/2, n < 0; 0, n = 0; 1/2, n > 0
1/2
...
−3 −2 −1
0 1 2 3
n
...
−1/2

Figure 1.18 Example of the even-odd decomposition of a discrete-time signal.

An important fact is that any signal can be broken into a sum of two signals, one of which is even and one of which is odd. To see this, consider the signal

$$\mathcal{E}v\{x(t)\} = \frac{1}{2}[x(t) + x(-t)], \tag{1.18}$$

which is referred to as the *even part* of $x(t)$. Similarly, the *odd part* of $x(t)$ is given by

$$\mathcal{O}d\{x(t)\} = \frac{1}{2}[x(t) - x(-t)]. \tag{1.19}$$

It is a simple exercise to check that the even part is in fact even, that the odd part is odd, and that $x(t)$ is the sum of the two. Exactly analogous definitions hold in the discrete-time case. An example of the even-odd decomposition of a discrete-time signal is given in Figure 1.18.

1.3 EXPONENTIAL AND SINUSOIDAL SIGNALS

In this section and the next, we introduce several basic continuous-time and discrete-time signals. Not only do these signals occur frequently, but they also serve as basic building blocks from which we can construct many other signals.

1.3.1 Continuous-Time Complex Exponential and Sinusoidal Signals

The continuous-time *complex exponential signal* is of the form

$$x(t) = Ce^{at}, \tag{1.20}$$

where C and a are, in general, complex numbers. Depending upon the values of these parameters, the complex exponential can exhibit several different characteristics.

Real Exponential Signals

As illustrated in Figure 1.19, if C and a are real [in which case $x(t)$ is called a *real exponential*], there are basically two types of behavior. If a is positive, then as t increases $x(t)$ is a growing exponential, a form that is used in describing many different physical processes, including chain reactions in atomic explosions and complex chemical reactions. If a is negative, then $x(t)$ is a decaying exponential, a signal that is also used to describe a wide variety of phenomena, including the process of radioactive decay and the responses of RC circuits and damped mechanical systems. In particular, as shown in Problems 2.61 and 2.62, the natural responses of the circuit in Figure 1.1 and the automobile in Figure 1.2 are decaying exponentials. Also, we note that for $a = 0$, $x(t)$ is constant.

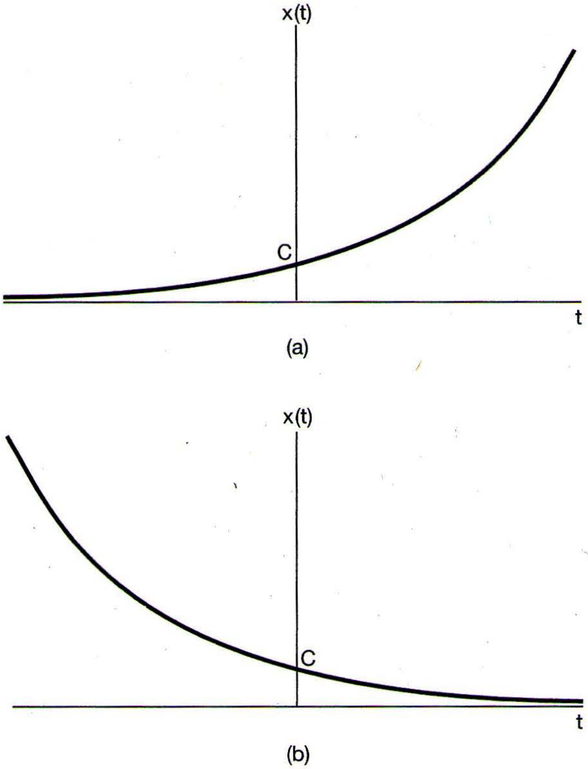


Figure 1.19 Continuous-time real exponential $x(t) = Ce^{at}$: (a) $a > 0$; (b) $a < 0$.

Periodic Complex Exponential and Sinusoidal Signals

A second important class of complex exponentials is obtained by constraining a to be purely imaginary. Specifically, consider

$$x(t) = e^{j\omega_0 t}. \tag{1.21}$$

An important property of this signal is that it is periodic. To verify this, we recall from eq. (1.11) that $x(t)$ will be periodic with period T if

$$e^{j\omega_0 t} = e^{j\omega_0(t+T)}. \tag{1.22}$$

Or, since

$$e^{j\omega_0(t+T)} = e^{j\omega_0 t}e^{j\omega_0 T},$$

it follows that for periodicity, we must have

$$e^{j\omega_0 T} = 1. \tag{1.23}$$

If $\omega_0 = 0$, then $x(t) = 1$, which is periodic for any value of T. If $\omega_0 \neq 0$, then the fundamental period T_0 of $x(t)$—that is, the smallest positive value of T for which eq. (1.23) holds—is

$$T_0 = \frac{2\pi}{|\omega_0|}. \tag{1.24}$$

Thus, the signals $e^{j\omega_0 t}$ and $e^{-j\omega_0 t}$ have the same fundamental period.

A signal closely related to the periodic complex exponential is the *sinusoidal signal*

$$x(t) = A\cos(\omega_0 t + \phi), \tag{1.25}$$

as illustrated in Figure 1.20. With seconds as the units of t, the units of ϕ and ω_0 are radians and radians per second, respectively. It is also common to write $\omega_0 = 2\pi f_0$, where f_0 has the units of cycles per second, or hertz (Hz). Like the complex exponential signal, the sinusoidal signal is periodic with fundamental period T_0 given by eq. (1.24). Sinusoidal and

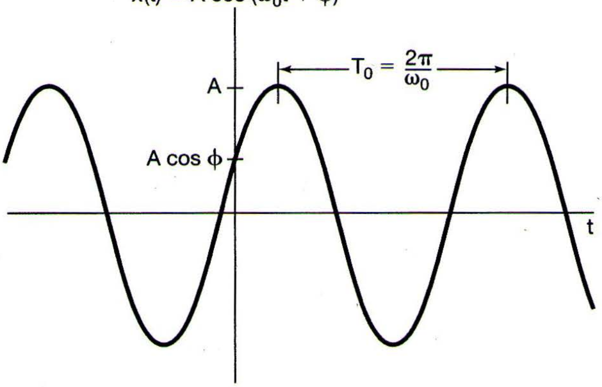


Figure 1.20 Continuous-time sinusoidal signal.

complex exponential signals are also used to describe the characteristics of many physical processes—in particular, physical systems in which energy is conserved. For example, as shown in Problem 2.61, the natural response of an *LC* circuit is sinusoidal, as is the simple harmonic motion of a mechanical system consisting of a mass connected by a spring to a stationary support. The acoustic pressure variations corresponding to a single musical tone are also sinusoidal.

By using Euler's relation,[2] the complex exponential in eq. (1.21) can be written in terms of sinusoidal signals with the same fundamental period:

$$e^{j\omega_0 t} = \cos\omega_0 t + j\sin\omega_0 t. \tag{1.26}$$

Similarly, the sinusoidal signal of eq. (1.25) can be written in terms of periodic complex exponentials, again with the same fundamental period:

$$A\cos(\omega_0 t + \phi) = \frac{A}{2}e^{j\phi}e^{j\omega_0 t} + \frac{A}{2}e^{-j\phi}e^{-j\omega_0 t}. \tag{1.27}$$

Note that the two exponentials in eq. (1.27) have complex amplitudes. Alternatively, we can express a sinusoid in terms of a complex exponential signal as

$$A\cos(\omega_0 t + \phi) = A\,\mathcal{R}e\{e^{j(\omega_0 t + \phi)}\}, \tag{1.28}$$

where, if c is a complex number, $\mathcal{R}e\{c\}$ denotes its real part. We will also use the notation $\mathcal{I}m\{c\}$ for the imaginary part of c, so that, for example,

$$A\sin(\omega_0 t + \phi) = A\,\mathcal{I}m\{e^{j(\omega_0 t + \phi)}\}. \tag{1.29}$$

From eq. (1.24), we see that the fundamental period T_0 of a continuous-time sinusoidal signal or a periodic complex exponential is inversely proportional to $|\omega_0|$, which we will refer to as the *fundamental frequency*. From Figure 1.21, we see graphically what this means. If we decrease the magnitude of ω_0, we slow down the rate of oscillation and therefore increase the period. Exactly the opposite effects occur if we increase the magnitude of ω_0. Consider now the case $\omega_0 = 0$. In this case, as we mentioned earlier, $x(t)$ is constant and therefore is periodic with period T for any positive value of T. Thus, the fundamental period of a constant signal is undefined. On the other hand, there is no ambiguity in defining the fundamental frequency of a constant signal to be zero. That is, a constant signal has a zero rate of oscillation.

Periodic signals—and in particular, the complex periodic exponential signal in eq. (1.21) and the sinusoidal signal in eq. (1.25)—provide important examples of signals with infinite total energy but finite average power. For example, consider the periodic exponential signal of eq. (1.21), and suppose that we calculate the total energy and average power in this signal over one period:

$$\begin{aligned} E_{\text{period}} &= \int_0^{T_0} \left|e^{j\omega_0 t}\right|^2 dt \\ &= \int_0^{T_0} 1 \cdot dt = T_0, \end{aligned} \tag{1.30}$$

---

[2] Euler's relation and other basic ideas related to the manipulation of complex numbers and exponentials are considered in the mathematical review section of the problems at the end of the chapter.

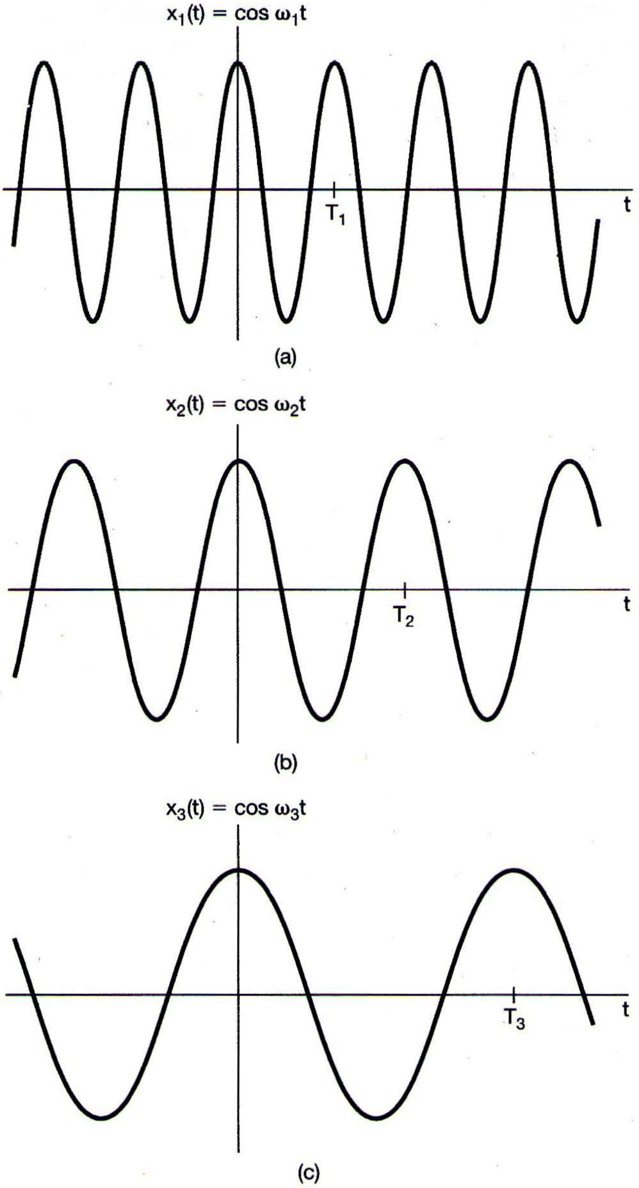


Figure 1.21 Relationship between the fundamental frequency and period for continuous-time sinusoidal signals; here, $\omega_1 > \omega_2 > \omega_3$, which implies that $T_1 < T_2 < T_3$.

$$P_{\text{period}} = \frac{1}{T_0} E_{\text{period}} = 1. \tag{1.31}$$

Since there are an infinite number of periods as t ranges from $-\infty$ to $+\infty$, the total energy integrated over all time is infinite. However, each period of the signal looks exactly the same. Since the average power of the signal equals 1 over each period, averaging over multiple periods always yields an average power of 1. That is, the complex periodic exponential signal has finite average power equal to

$$P_\infty = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \left|e^{j\omega_0 t}\right|^2 dt = 1. \tag{1.32}$$

Problem 1.3 provides additional examples of energy and power calculations for periodic and aperiodic signals.

Periodic complex exponentials will play a central role in much of our treatment of signals and systems, in part because they serve as extremely useful building blocks for many other signals. We will often find it useful to consider sets of *harmonically related* complex exponentials—that is, sets of periodic exponentials, all of which are periodic with a common period T_0. Specifically, a necessary condition for a complex exponential $e^{j\omega t}$ to be periodic with period T_0 is that

$$e^{j\omega T_0} = 1, \tag{1.33}$$

which implies that ωT_0 is a multiple of 2π, i.e.,

$$\omega T_0 = 2\pi k, \qquad k = 0, \pm 1, \pm 2, \ldots. \tag{1.34}$$

Thus, if we define

$$\omega_0 = \frac{2\pi}{T_0}, \tag{1.35}$$

we see that, to satisfy eq. (1.34), ω must be an integer multiple of ω_0. That is, a harmonically related set of complex exponentials is a set of periodic exponentials with fundamental frequencies that are all multiples of a single positive frequency ω_0:

$$\phi_k(t) = e^{jk\omega_0 t}, \qquad k = 0, \pm 1, \pm 2, \ldots. \tag{1.36}$$

For $k = 0$, $\phi_k(t)$ is a constant, while for any other value of k, $\phi_k(t)$ is periodic with fundamental frequency $|k|\omega_0$ and fundamental period

$$\frac{2\pi}{|k|\omega_0} = \frac{T_0}{|k|}. \tag{1.37}$$

The kth harmonic $\phi_k(t)$ is still periodic with period T_0 as well, as it goes through exactly $|k|$ of its fundamental periods during any time interval of length T_0.

Our use of the term "harmonic" is consistent with its use in music, where it refers to tones resulting from variations in acoustic pressure at frequencies that are integer multiples of a fundamental frequency. For example, the pattern of vibrations of a string on an instrument such as a violin can be described as a superposition—i.e., a weighted sum—of harmonically related periodic exponentials. In Chapter 3, we will see that we can build a very rich class of periodic signals using the harmonically related signals of eq. (1.36) as the building blocks.

Example 1.5

It is sometimes desirable to express the sum of two complex exponentials as the product of a single complex exponential and a single sinusoid. For example, suppose we wish to plot the magnitude of the signal

$$x(t) = e^{j2t} + e^{j3t}. \tag{1.38}$$

To do this, we first factor out a complex exponential from the right side of eq. (1.38), where the frequency of this exponential factor is taken as the average of the frequencies of the two exponentials in the sum. Doing this, we obtain

$$x(t) = e^{j2.5t}(e^{-j0.5} + e^{j0.5t}), \tag{1.39}$$

which, because of Euler's relation, can be rewritten as

$$x(t) = 2e^{j2.5t}\cos(0.5t). \tag{1.40}$$

From this, we can directly obtain an expression for the magnitude of $x(t)$:

$$|x(t)| = 2|\cos(0.5t)|. \tag{1.41}$$

Here, we have used the fact that the magnitude of the complex exponential $e^{j2.5t}$ is always unity. Thus, $|x(t)|$ is what is commonly referred to as a full-wave rectified sinusoid, as shown in Figure 1.22.

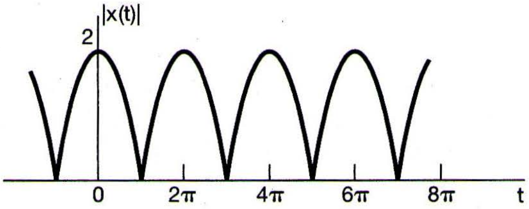


Figure 1.22 The full-wave rectified sinusoid of Example 1.5.

General Complex Exponential Signals

The most general case of a complex exponential can be expressed and interpreted in terms of the two cases we have examined so far: the real exponential and the periodic complex exponential. Specifically, consider a complex exponential Ce^{at}, where C is expressed in polar form and a in rectangular form. That is,

$$C = |C|e^{j\theta}$$

and

$$a = r + j\omega_0.$$

Then

$$Ce^{at} = |C|e^{j\theta}e^{(r+j\omega_0)t} = |C|e^{rt}e^{j(\omega_0 t+\theta)}. \tag{1.42}$$

Using Euler's relation, we can expand this further as

$$Ce^{at} = |C|e^{rt}\cos(\omega_0 t + \theta) + j|C|e^{rt}\sin(\omega_0 t + \theta). \tag{1.43}$$

Thus, for $r = 0$, the real and imaginary parts of a complex exponential are sinusoidal. For $r > 0$ they correspond to sinusoidal signals multiplied by a growing exponential, and for $r < 0$ they correspond to sinusoidal signals multiplied by a decaying exponential. These two cases are shown in Figure 1.23. The dashed lines in the figure correspond to the functions $\pm|C|e^{rt}$. From eq. (1.42), we see that $|C|e^{rt}$ is the magnitude of the complex exponential. Thus, the dashed curves act as an envelope for the oscillatory curve in the figure in that the peaks of the oscillations just reach these curves, and in this way the envelope provides us with a convenient way to visualize the general trend in the amplitude of the oscillations.

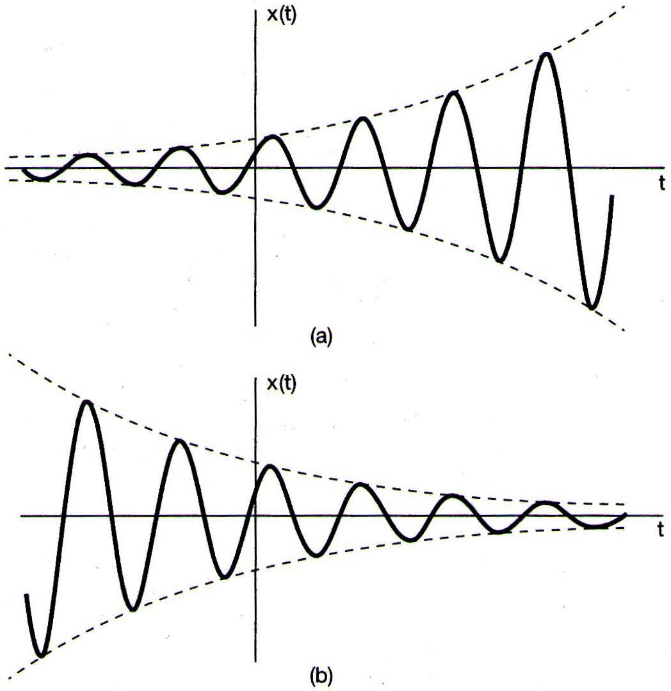


Figure 1.23 (a) Growing sinusoidal signal $x(t) = Ce^{rt}\cos(\omega_0 t + \theta)$, $r > 0$; (b) decaying sinusoid $x(t) = Ce^{rt}\cos(\omega_0 t + \theta)$, $r < 0$.

Sinusoidal signals multiplied by decaying exponentials are commonly referred to as *damped sinusoids.* Examples of damped sinusoids arise in the response of *RLC* circuits and in mechanical systems containing both damping and restoring forces, such as automotive suspension systems. These kinds of systems have mechanisms that dissipate energy (resistors, damping forces such as friction) with oscillations that decay in time. Examples illustrating such systems and their damped sinusoidal natural responses can be found in Problems 2.61 and 2.62.

1.3.2 Discrete-Time Complex Exponential and Sinusoidal Signals

As in continuous time, an important signal in discrete time is the *complex exponential signal* or *sequence,* defined by

$$x[n] = C\alpha^n, \tag{1.44}$$

where C and α are, in general, complex numbers. This could alternatively be expressed in the form

$$x[n] = Ce^{\beta n}, \tag{1.45}$$

where

$$\alpha = e^{\beta}.$$

Although the form of the discrete-time complex exponential sequence given in eq. (1.45) is more analogous to the form of the continuous-time exponential, it is often more convenient to express the discrete-time complex exponential sequence in the form of eq. (1.44).

Real Exponential Signals

If C and α are real, we can have one of several types of behavior, as illustrated in Figure 1.24. If $|\alpha| > 1$ the magnitude of the signal grows exponentially with n, while if $|\alpha| < 1$ we have a decaying exponential. Furthermore, if α is positive, all the values of $C\alpha^n$ are of the same sign, but if α is negative then the sign of $x[n]$ alternates. Note also that if $\alpha = 1$ then $x[n]$ is a constant, whereas if $\alpha = -1$, $x[n]$ alternates in value between $+C$ and $-C$. Real-valued discrete-time exponentials are often used to describe population growth as a function of generation and total return on investment as a function of day, month, or quarter.

Sinusoidal Signals

Another important complex exponential is obtained by using the form given in eq. (1.45) and by constraining β to be purely imaginary (so that $|\alpha| = 1$). Specifically, consider

$$x[n] = e^{j\omega_0 n}. \tag{1.46}$$

As in the continuous-time case, this signal is closely related to the sinusoidal signal

$$x[n] = A\cos(\omega_0 n + \phi). \tag{1.47}$$

If we take n to be dimensionless, then both ω_0 and ϕ have units of radians. Three examples of sinusoidal sequences are shown in Figure 1.25.

As before, Euler's relation allows us to relate complex exponentials and sinusoids:

$$e^{j\omega_0 n} = \cos\omega_0 n + j\sin\omega_0 n \tag{1.48}$$

and

$$A\cos(\omega_0 n + \phi) = \frac{A}{2}e^{j\phi}e^{j\omega_0 n} + \frac{A}{2}e^{-j\phi}e^{-j\omega_0 n}. \tag{1.49}$$

The signals in eqs. (1.46) and (1.47) are examples of discrete-time signals with infinite total energy but finite average power. For example, since $|e^{j\omega_0 n}|^2 = 1$, every sample of the signal in eq. (1.46) contributes 1 to the signal's energy. Thus, the total energy for $-\infty < n < \infty$ is infinite, while the average power per time point is obviously equal to 1. Other examples of energy and power calculations for discrete-time signals are given in Problem 1.3.

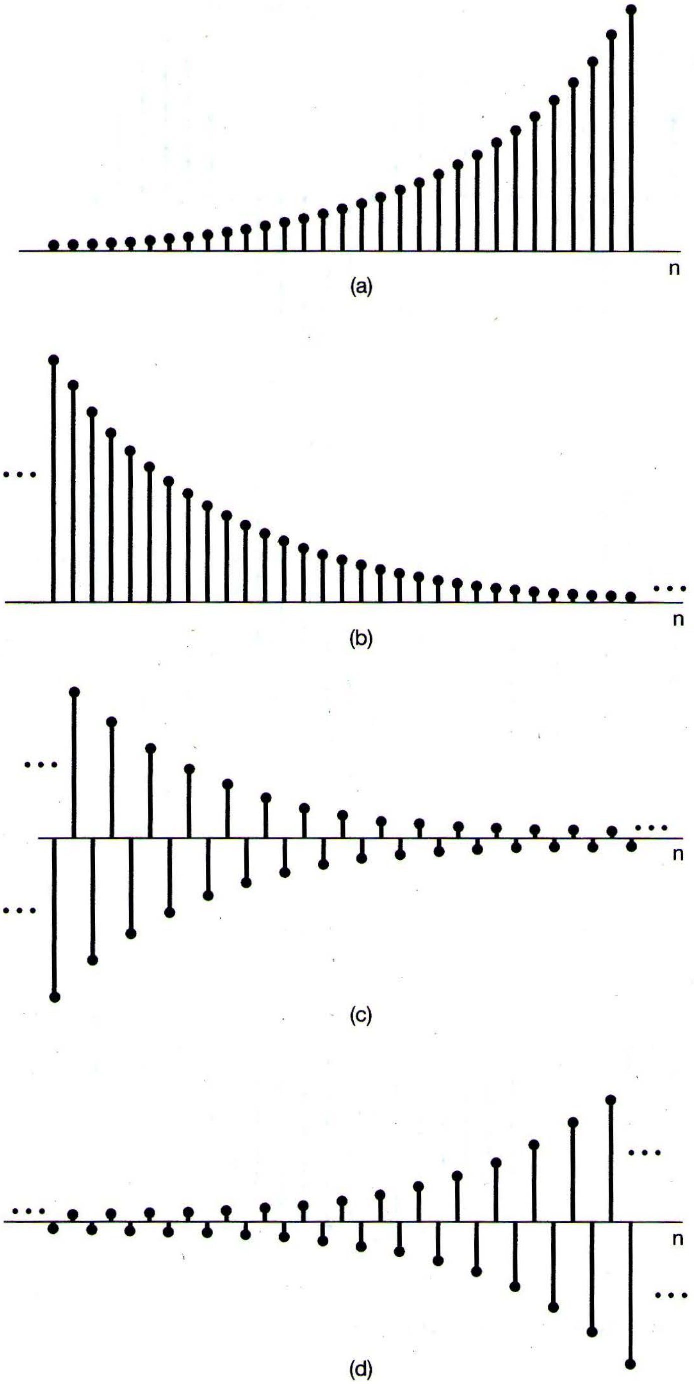
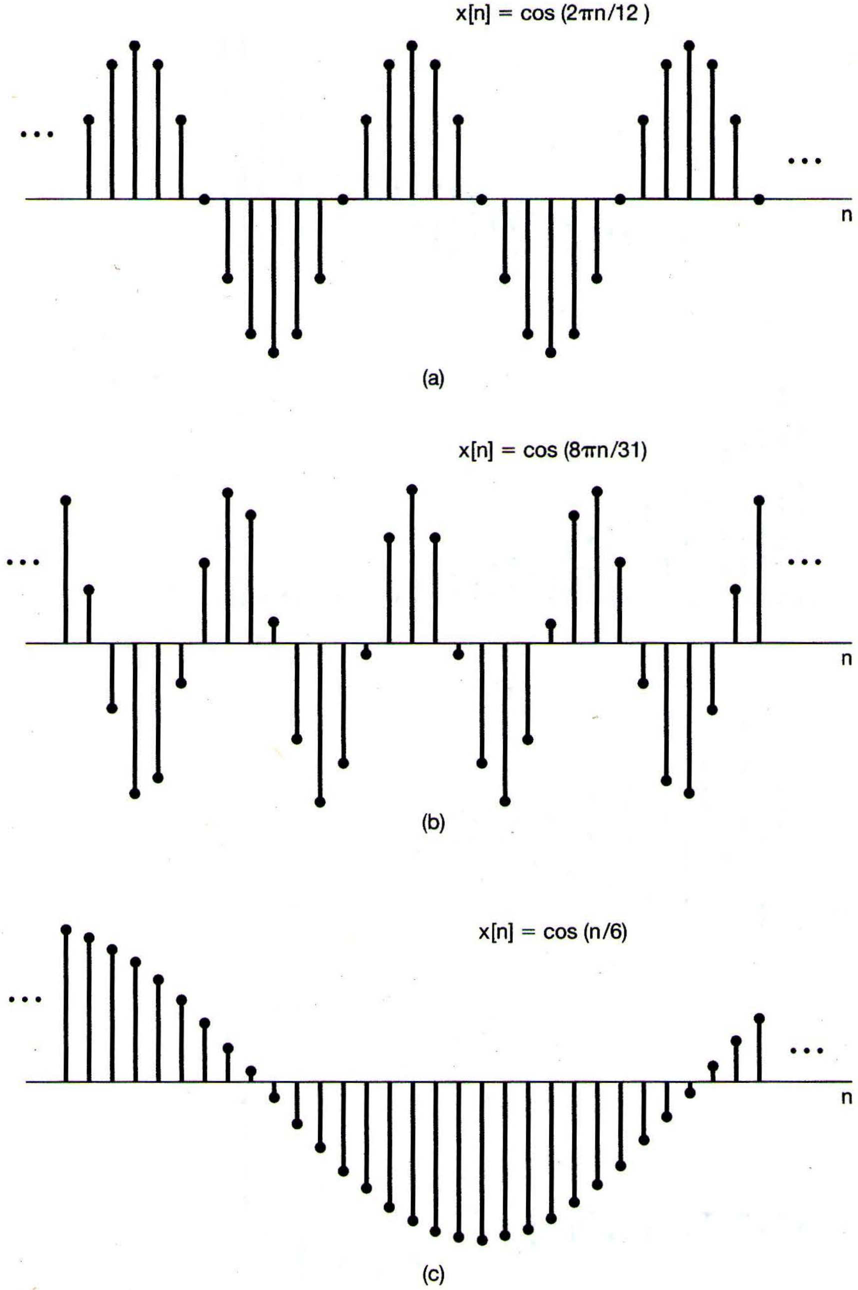


Figure 1.24 The real exponential signal $x[n] = C\alpha^n$:
(a) $\alpha > 1$; (b) $0 < \alpha < 1$;
(c) $-1 < \alpha < 0$; (d) $\alpha < -1$.


$x[n] = \cos(2\pi n/12)$

(a)

$x[n] = \cos(8\pi n/31)$

(b)

$x[n] = \cos(n/6)$

(c)

Figure 1.25 Discrete-time sinusoidal signals.

General Complex Exponential Signals

The general discrete-time complex exponential can be written and interpreted in terms of real exponentials and sinusoidal signals. Specifically, if we write C and α in polar form, viz.,

$$C = |C|e^{j\theta}$$

and

$$\alpha = |\alpha|e^{j\omega_0},$$

then

$$C\alpha^n = |C||\alpha|^n \cos(\omega_0 n + \theta) + j|C||\alpha|^n \sin(\omega_0 n + \theta). \tag{1.50}$$

Thus, for $|\alpha| = 1$, the real and imaginary parts of a complex exponential sequence are sinusoidal. For $|\alpha| < 1$ they correspond to sinusoidal sequences multiplied by a decaying exponential, while for $|\alpha| > 1$ they correspond to sinusoidal sequences multiplied by a growing exponential. Examples of these signals are depicted in Figure 1.26.

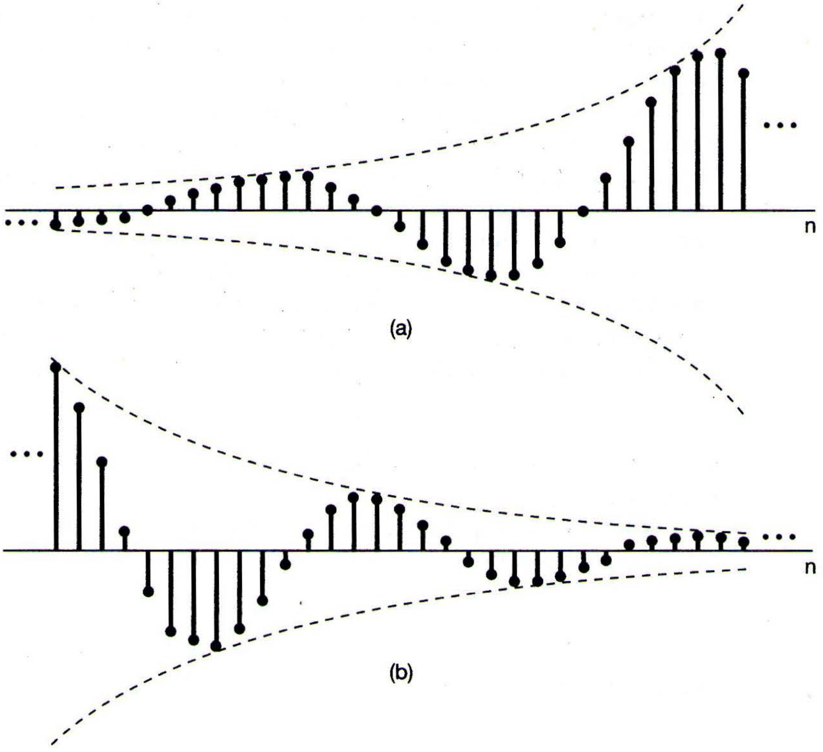


Figure 1.26 (a) Growing discrete-time sinusoidal signals; (b) decaying discrete-time sinusoid.

1.3.3 Periodicity Properties of Discrete-Time Complex Exponentials

While there are many similarities between continuous-time and discrete-time signals, there are also a number of important differences. One of these concerns the discrete-time exponential signal $e^{j\omega_0 n}$. In Section 1.3.1, we identified the following two properties of its continuous-time counterpart $e^{j\omega_0 t}$: (1) the larger the magnitude of ω_0, the higher is the rate of oscillation in the signal; and (2) $e^{j\omega_0 t}$ is periodic for any value of ω_0. In this section we describe the discrete-time versions of both of these properties, and as we will see, there are definite differences between each of these and its continuous-time counterpart.

The fact that the first of these properties is different in discrete time is a direct consequence of another extremely important distinction between discrete-time and continuous-time complex exponentials. Specifically, consider the discrete-time complex exponential with frequency $\omega_0 + 2\pi$:

$$e^{j(\omega_0+2\pi)n} = e^{j2\pi n}e^{j\omega_0 n} = e^{j\omega_0 n}. \tag{1.51}$$

From eq. (1.51), we see that the exponential at frequency $\omega_0 + 2\pi$ is the *same* as that at frequency ω_0. Thus, we have a very different situation from the continuous-time case, in which the signals $e^{j\omega_0 t}$ are all distinct for distinct values of ω_0. In discrete time, these signals are not distinct, as the signal with frequency ω_0 is identical to the signals with frequencies $\omega_0 \pm 2\pi$, $\omega_0 \pm 4\pi$, and so on. Therefore, in considering discrete-time complex exponentials, we need only consider a frequency interval of length 2π in which to choose ω_0. Although, according to eq. (1.51), any interval of length 2π will do, on most occasions we will use the interval $0 \leq \omega_0 < 2\pi$ or the interval $-\pi \leq \omega_0 < \pi$.

Because of the periodicity implied by eq. (1.51), the signal $e^{j\omega_0 n}$ does *not* have a continually increasing rate of oscillation as ω_0 is increased in magnitude. Rather, as illustrated in Figure 1.27, as we increase ω_0 from 0, we obtain signals that oscillate more and more rapidly until we reach $\omega_0 = \pi$. As we continue to increase ω_0, we *decrease* the rate of oscillation until we reach $\omega_0 = 2\pi$, which produces the same constant sequence as $\omega_0 = 0$. Therefore, the low-frequency (that is, slowly varying) discrete-time exponentials have values of ω_0 near 0, 2π, and any other even multiple of π, while the high frequencies (corresponding to rapid variations) are located near $\omega_0 = \pm\pi$ and other odd multiples of π. Note in particular that for $\omega_0 = \pi$ or any other odd multiple of π,

$$e^{j\pi n} = (e^{j\pi})^n = (-1)^n, \tag{1.52}$$

so that this signal oscillates rapidly, changing sign at each point in time [as illustrated in Figure 1.27(e)].

The second property we wish to consider concerns the periodicity of the discrete-time complex exponential. In order for the signal $e^{j\omega_0 n}$ to be periodic with period $N > 0$, we must have

$$e^{j\omega_0(n+N)} = e^{j\omega_0 n}, \tag{1.53}$$

or equivalently,

$$e^{j\omega_0 N} = 1. \tag{1.54}$$

For eq. (1.54) to hold, $\omega_0 N$ must be a multiple of 2π. That is, there must be an integer m such that

$$\omega_0 N = 2\pi m, \tag{1.55}$$

or equivalently,

$$\frac{\omega_0}{2\pi} = \frac{m}{N}. \tag{1.56}$$

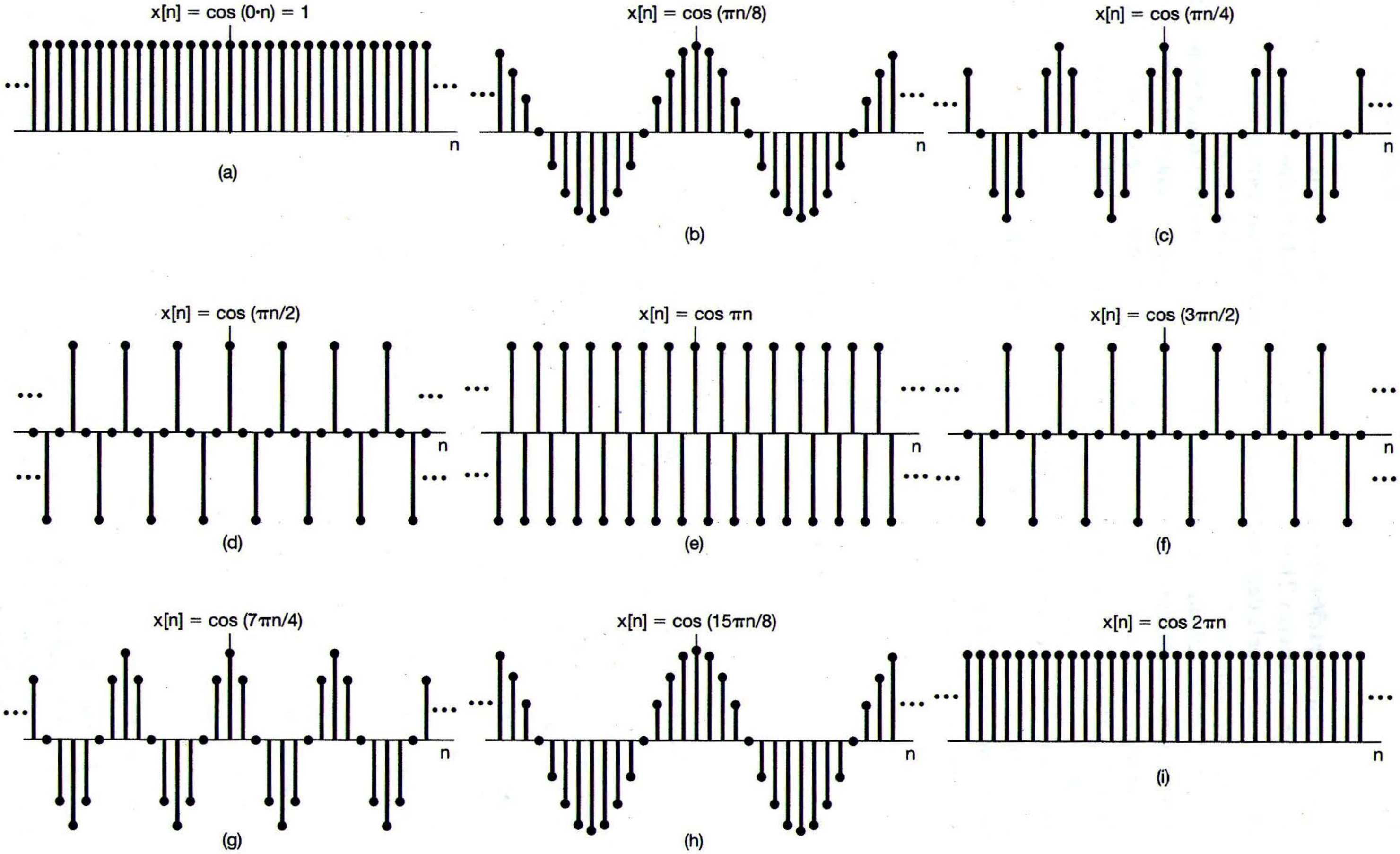


Figure 1.27 Discrete-time sinusoidal sequences for several different frequencies.

According to eq. (1.56), the signal $e^{j\omega_0 n}$ is periodic if $\omega_0/2\pi$ is a rational number and is not periodic otherwise. These same observations also hold for discrete-time sinusoids. For example, the signals depicted in Figure 1.25(a) and (b) are periodic, while the signal in Figure 1.25(c) is not.

Using the calculations that we have just made, we can also determine the fundamental period and frequency of discrete-time complex exponentials, where we define the fundamental frequency of a discrete-time periodic signal as we did in continuous time. That is, if $x[n]$ is periodic with fundamental period N, its fundamental frequency is $2\pi/N$. Consider, then, a periodic complex exponential $x[n] = e^{j\omega_0 n}$ with $\omega_0 \neq 0$. As we have just seen, ω_0 must satisfy eq. (1.56) for some pair of integers m and N, with $N > 0$. In Problem 1.35, it is shown that if $\omega_0 \neq 0$ and if N and m have no factors in common, then the fundamental period of $x[n]$ is N. Using this fact together with eq. (1.56), we find that the fundamental frequency of the periodic signal $e^{j\omega_0 n}$ is

$$\frac{2\pi}{N} = \frac{\omega_0}{m}. \tag{1.57}$$

Note that the fundamental period can also be written as

$$N = m\left(\frac{2\pi}{\omega_0}\right). \tag{1.58}$$

These last two expressions again differ from their continuous-time counterparts. In Table 1.1, we have summarized some of the differences between the continuous-time signal $e^{j\omega_0 t}$ and the discrete-time signal $e^{j\omega_0 n}$. Note that, as in the continuous-time case, the constant discrete-time signal resulting from setting $\omega_0 = 0$ has a fundamental frequency of zero, and its fundamental period is undefined.

TABLE 1.1 Comparison of the signals $e^{j\omega_0 t}$ and $e^{j\omega_0 n}$.

$e^{j\omega_0 t}$	$e^{j\omega_0 n}$
Distinct signals for distinct values of ω_0	Identical signals for values of ω_0 separated by multiples of 2π
Periodic for any choice of ω_0	Periodic only if $\omega_0 = 2\pi m/N$ for some integers $N > 0$ and m.
Fundamental frequency ω_0	Fundamental frequency* ω_0/m
Fundamental period $\omega_0 = 0$: undefined $\omega_0 \neq 0$: $\frac{2\pi}{\omega_0}$	Fundamental period* $\omega_0 = 0$: undefined $\omega_0 \neq 0$: $m\left(\frac{2\pi}{\omega_0}\right)$

*Assumes that m and N do not have any factors in common.

To gain some additional insight into these properties, let us examine again the signals depicted in Figure 1.25. First, consider the sequence $x[n] = \cos(2\pi n/12)$, depicted in Figure 1.25(a), which we can think of as the set of samples of the continuous-time sinusoid $x(t) = \cos(2\pi t/12)$ at integer time points. In this case, $x(t)$ is periodic with fundamental period 12 and $x[n]$ is also periodic with fundamental period 12. That is, the values of $x[n]$ repeat every 12 points, exactly in step with the fundamental period of $x(t)$.

In contrast, consider the signal $x[n] = \cos(8\pi n/31)$, depicted in Figure 1.25(b), which we can view as the set of samples of $x(t) = \cos(8\pi t/31)$ at integer points in time. In this case, $x(t)$ is periodic with fundamental period 31/4. On the other hand, $x[n]$ is periodic with fundamental period 31. The reason for this difference is that the discrete-time signal is defined only for integer values of the independent variable. Thus, there is no sample at time $t = 31/4$, when $x(t)$ completes one period (starting from $t = 0$). Similarly, there is no sample at $t = 2 \cdot 31/4$ or $t = 3 \cdot 31/4$, when $x(t)$ has completed two or three periods, but there is a sample at $t = 4 \cdot 31/4 = 31$, when $x(t)$ has completed *four* periods. This can be seen in Figure 1.25(b), where the pattern of $x[n]$ values does *not* repeat with each single cycle of positive and negative values. Rather, the pattern repeats after four such cycles, namely, every 31 points.

Similarly, the signal $x[n] = \cos(n/6)$ can be viewed as the set of samples of the signal $x(t) = \cos(t/6)$ at integer time points. In this case, the values of $x(t)$ at integer sample points never repeat, as these sample points never span an interval that is an exact multiple of the period, 12π, of $x(t)$. Thus, $x[n]$ is not periodic, although the eye visually interpolates between the sample points, suggesting the envelope $x(t)$, which *is* periodic. The use of the concept of sampling to gain insight into the periodicity of discrete-time sinusoidal sequences is explored further in Problem 1.36.

Example 1.6

Suppose that we wish to determine the fundamental period of the discrete-time signal

$$x[n] = e^{j(2\pi/3)n} + e^{j(3\pi/4)n}. \tag{1.59}$$

The first exponential on the right-hand side of eq. (1.59) has a fundamental period of 3. While this can be verified from eq. (1.58), there is a simpler way to obtain that answer. In particular, note that the angle $(2\pi/3)n$ of the first term must be incremented by a multiple of 2π for the values of this exponential to begin repeating. We then immediately see that if n is incremented by 3, the angle will be incremented by a single multiple of 2π. With regard to the second term, we see that incrementing the angle $(3\pi/4)n$ by 2π would require n to be incremented by 8/3, which is impossible, since n is restricted to being an integer. Similarly, incrementing the angle by 4π would require a noninteger increment of 16/3 to n. However, incrementing the angle by 6π requires an increment of 8 to n, and thus the fundamental period of the second term is 8.

Now, for the entire signal $x[n]$ to repeat, each of the terms in eq. (1.59) must go through an integer number of its own fundamental period. The smallest increment of n that accomplishes this is 24. That is, over an interval of 24 points, the first term on the right-hand side of eq. (1.59) will have gone through eight of its fundamental periods, the second term through three of its fundamental periods, and the overall signal $x[n]$ through exactly one of its fundamental periods.

As in continuous time, it is also of considerable value in discrete-time signal and system analysis to consider sets of harmonically related periodic exponentials—that is, periodic exponentials with a common period N. From eq. (1.56), we know that these are precisely the signals which are at frequencies which are multiples of $2\pi/N$. That is,

$$\phi_k[n] = e^{jk(2\pi/N)n}, \qquad k = 0, \pm 1, \ldots. \tag{1.60}$$

In the continuous-time case, all of the harmonically related complex exponentials $e^{jk(2\pi/T)t}$, $k = 0, \pm 1, \pm 2, \ldots$, are distinct. However, because of eq. (1.51), this is *not* the case in discrete time. Specifically,

$$\begin{aligned}\phi_{k+N}[n] &= e^{j(k+N)(2\pi/N)n} \\ &= e^{jk(2\pi/N)n}e^{j2\pi n} = \phi_k[n].\end{aligned} \tag{1.61}$$

This implies that there are only N distinct periodic exponentials in the set given in eq. (1.60). For example,

$$\phi_0[n] = 1, \phi_1[n] = e^{j2\pi n/N}, \phi_2[n] = e^{j4\pi n/N}, \ldots, \phi_{N-1}[n] = e^{j2\pi(N-1)n/N} \tag{1.62}$$

are all distinct, and any other $\phi_k[n]$ is identical to one of these (e.g., $\phi_N[n] = \phi_0[n]$ and $\phi_{-1}[n] = \phi_{N-1}[n]$).

1.4 THE UNIT IMPULSE AND UNIT STEP FUNCTIONS

In this section, we introduce several other basic signals—specifically, the unit impulse and step functions in continuous and discrete time—that are also of considerable importance in signal and system analysis. In Chapter 2, we will see how we can use unit impulse signals as basic building blocks for the construction and representation of other signals. We begin with the discrete-time case.

1.4.1 The Discrete-Time Unit Impulse and Unit Step Sequences

One of the simplest discrete-time signals is the *unit impulse* (or *unit sample*), which is defined as

$$\delta[n] = \begin{cases} 0, & n \neq 0 \\ 1, & n = 0 \end{cases} \tag{1.63}$$

and which is shown in Figure 1.28. Throughout the book, we will refer to $\delta[n]$ interchangeably as the unit impulse or unit sample.

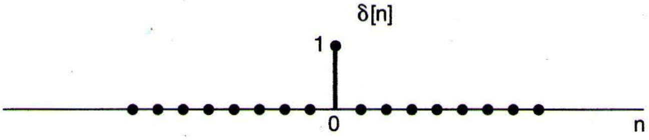


Figure 1.28 Discrete-time unit impulse (sample).

A second basic discrete-time signal is the discrete-time *unit step,* denoted by $u[n]$ and defined by

$$u[n] = \begin{cases} 0, & n < 0 \\ 1, & n \geq 0 \end{cases}. \tag{1.64}$$

The unit step sequence is shown in Figure 1.29.

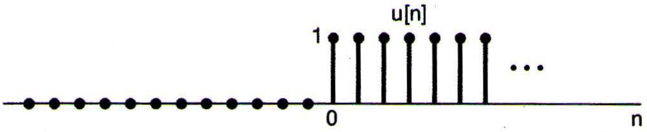
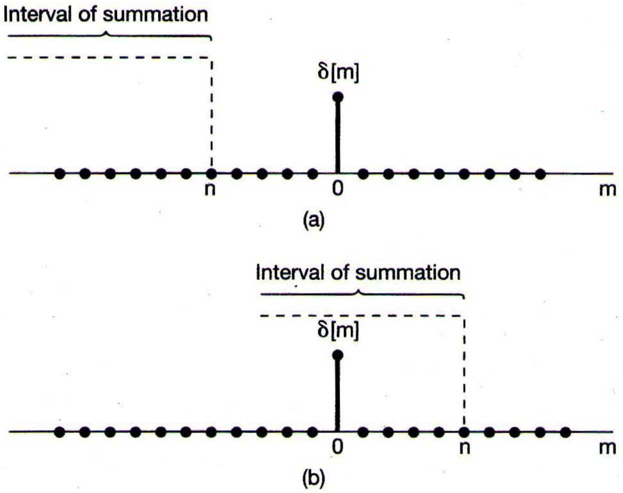


Figure 1.29 Discrete-time unit step sequence.



(a)

(b)

Figure 1.30 Running sum of eq. (1.66): (a) $n < 0$; (b) $n > 0$.

There is a close relationship between the discrete-time unit impulse and unit step. In particular, the discrete-time unit impulse is the *first difference* of the discrete-time step

$$\delta[n] = u[n] - u[n-1]. \tag{1.65}$$

Conversely, the discrete-time unit step is the *running sum* of the unit sample. That is,

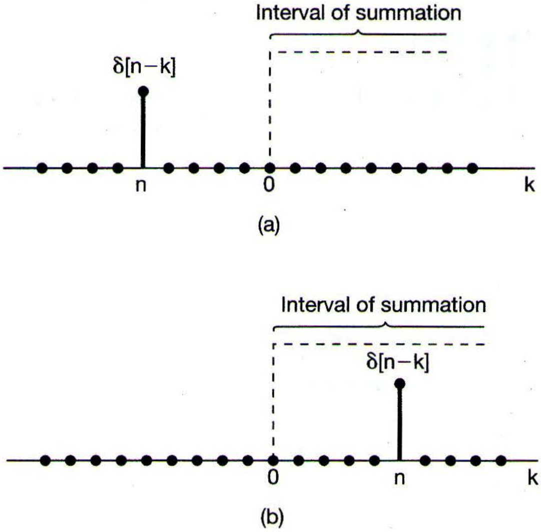
$$u[n] = \sum_{m=-\infty}^{n} \delta[m]. \tag{1.66}$$

Equation (1.66) is illustrated graphically in Figure 1.30. Since the only nonzero value of the unit sample is at the point at which its argument is zero, we see from the figure that the running sum in eq. (1.66) is 0 for $n < 0$ and 1 for $n \geq 0$. Furthermore, by changing the variable of summation from m to $k = n - m$ in eq. (1.66), we find that the discrete-time unit step can also be written in terms of the unit sample as

$$u[n] = \sum_{k=\infty}^{0} \delta[n-k],$$

or equivalently,

$$u[n] = \sum_{k=0}^{\infty} \delta[n-k]. \tag{1.67}$$



(a)

(b)

Figure 1.31 Relationship given in eq. (1.67): (a) $n < 0$; (b) $n > 0$.

Equation (1.67) is illustrated in Figure 1.31. In this case the nonzero value of $\delta[n-k]$ is at the value of k equal to n, so that again we see that the summation in eq. (1.67) is 0 for $n < 0$ and 1 for $n \geq 0$.

An interpretation of eq. (1.67) is as a superposition of delayed impulses; i.e., we can view the equation as the sum of a unit impulse $\delta[n]$ at $n = 0$, a unit impulse $\delta[n-1]$ at $n = 1$, another, $\delta[n-2]$, at $n = 2$, etc. We will make explicit use of this interpretation in Chapter 2.

The unit impulse sequence can be used to sample the value of a signal at $n = 0$. In particular, since $\delta[n]$ is nonzero (and equal to 1) only for $n = 0$, it follows that

$$x[n]\delta[n] = x[0]\delta[n]. \tag{1.68}$$

More generally, if we consider a unit impulse $\delta[n-n_0]$ at $n = n_0$, then

$$x[n]\delta[n-n_0] = x[n_0]\delta[n-n_0]. \tag{1.69}$$

This sampling property of the unit impulse will play an important role in Chapters 2 and 7.

1.4.2 The Continuous-Time Unit Step and Unit Impulse Functions

The continuous-time *unit step function* $u(t)$ is defined in a manner similar to its discrete-time counterpart. Specifically,

$$u(t) = \begin{cases} 0, & t < 0 \\ 1, & t > 0 \end{cases}, \tag{1.70}$$

as is shown in Figure 1.32. Note that the unit step is discontinuous at $t = 0$. The continuous-time *unit impulse function* $\delta(t)$ is related to the unit step in a manner analogous

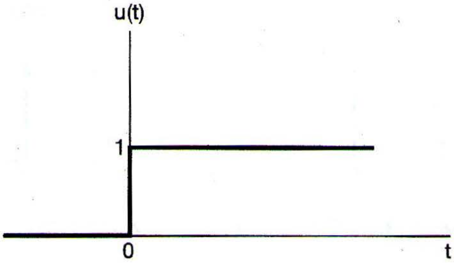


Figure 1.32 Continuous-time unit step function.

to the relationship between the discrete-time unit impulse and step functions. In particular, the continuous-time unit step is the *running integral* of the unit impulse

$$u(t) = \int_{-\infty}^{t} \delta(\tau)\, d\tau. \tag{1.71}$$

This also suggests a relationship between $\delta(t)$ and $u(t)$ analogous to the expression for $\delta[n]$ in eq. (1.65). In particular, it follows from eq. (1.71) that the continuous-time unit impulse can be thought of as the *first derivative* of the continuous-time unit step:

$$\delta(t) = \frac{du(t)}{dt}. \tag{1.72}$$

In contrast to the discrete-time case, there is some formal difficulty with this equation as a representation of the unit impulse function, since $u(t)$ is discontinuous at $t = 0$ and consequently is formally not differentiable. We can, however, interpret eq. (1.72) by considering an approximation to the unit step $u_\Delta(t)$, as illustrated in Figure 1.33, which rises from the value 0 to the value 1 in a short time interval of length Δ. The unit step, of course, changes values instantaneously and thus can be thought of as an idealization of $u_\Delta(t)$ for Δ so short that its duration doesn't matter for any practical purpose. Formally, $u(t)$ is the limit of $u_\Delta(t)$ as $\Delta \to 0$. Let us now consider the derivative

$$\delta_\Delta(t) = \frac{du_\Delta(t)}{dt}, \tag{1.73}$$

as shown in Figure 1.34.

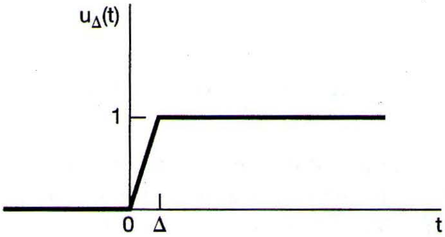


Figure 1.33 Continuous approximation to the unit step, $u_\Delta(t)$.

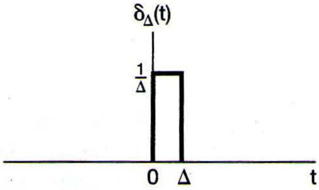


Figure 1.34 Derivative of $u_\Delta(t)$.

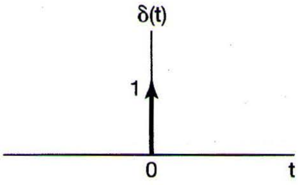


Figure 1.35 Continuous-time unit impulse.

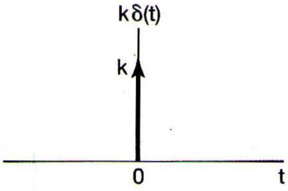


Figure 1.36 Scaled impulse.

Note that $\delta_\Delta(t)$ is a short pulse, of duration Δ and with unit area for any value of Δ. As $\Delta \rightarrow 0$, $\delta_\Delta(t)$ becomes narrower and higher, maintaining its unit area. Its limiting form,

$$\delta(t) = \lim_{\Delta\rightarrow 0} \delta_\Delta(t), \tag{1.74}$$

can then be thought of as an idealization of the short pulse $\delta_\Delta(t)$ as the duration Δ becomes insignificant. Since $\delta(t)$ has, in effect, no duration but unit area, we adopt the graphical notation for it shown in Figure 1.35, where the arrow at $t = 0$ indicates that the area of the pulse is concentrated at $t = 0$ and the height of the arrow and the "1" next to the arrow are used to represent the *area* of the impulse. More generally, a scaled impulse $k\delta(t)$ will have an area k, and thus,

$$\int_{-\infty}^{t} k\delta(\tau)\,d\tau = ku(t).$$

A scaled impulse with area k is shown in Figure 1.36, where the height of the arrow used to depict the scaled impulse is chosen to be proportional to the area of the impulse.

As with discrete time, we can provide a simple graphical interpretation of the running integral of eq. (1.71); this is shown in Figure 1.37. Since the area of the continuous-time unit impulse $\delta(\tau)$ is concentrated at $\tau = 0$, we see that the running integral is 0 for $t < 0$ and 1 for $t > 0$. Also, we note that the relationship in eq. (1.71) between the continuous-time unit step and impulse can be rewritten in a different form, analogous to the discrete-time form in eq. (1.67), by changing the variable of integration from τ to $\sigma = t - \tau$:

$$u(t) = \int_{-\infty}^{t} \delta(\tau)\,d\tau = \int_{\infty}^{0} \delta(t-\sigma)(-d\sigma),$$

or equivalently,

$$u(t) = \int_{0}^{\infty} \delta(t-\sigma)\,d\sigma. \tag{1.75}$$

The graphical interpretation of this form of the relationship between $u(t)$ and $\delta(t)$ is given in Figure 1.38. Since in this case the area of $\delta(t-\sigma)$ is concentrated at the point $\sigma = t$, we again see that the integral in eq. (1.75) is 0 for $t < 0$ and 1 for $t > 0$. This type of graphical interpretation of the behavior of the unit impulse under integration will be extremely useful in Chapter 2.

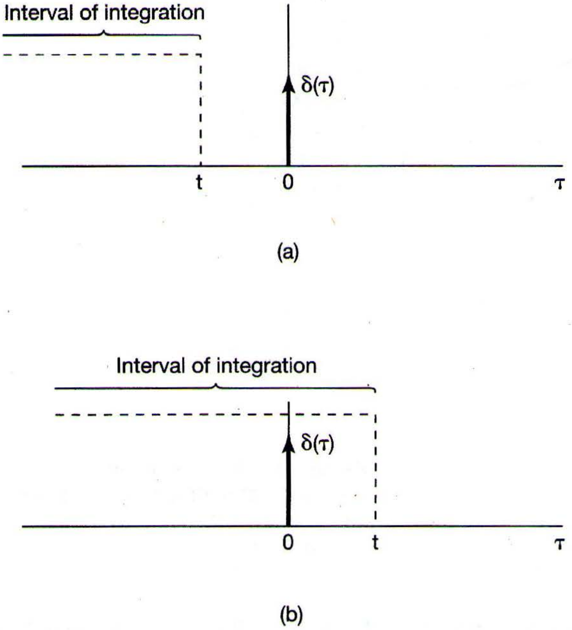

Interval of integration
δ(τ)
t
0
τ
(a)
Interval of integration
δ(τ)
0
t
τ
(b)

Figure 1.37 Running integral given in eq. (1.71): (a) $t < 0$; (b) $t > 0$.

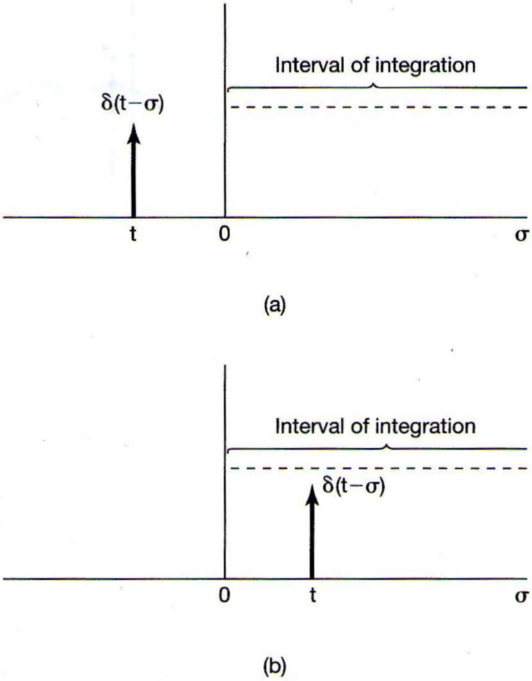

Interval of integration
δ(t−σ)
t
0
σ
(a)
Interval of integration
δ(t−σ)
0
t
σ
(b)

Figure 1.38 Relationship given in eq. (1.75): (a) $t < 0$; (b) $t > 0$.

As with the discrete-time impulse, the continuous-time impulse has a very important sampling property. In particular, for a number of reasons it will be important to consider the product of an impulse and more well-behaved continuous-time functions $x(t)$. The interpretation of this quantity is most readily developed using the definition of $\delta(t)$ according to eq. (1.74). Specifically, consider

$$x_1(t) = x(t)\delta_\Delta(t).$$

In Figure 1.39(a) we have depicted the two time functions $x(t)$ and $\delta_\Delta(t)$, and in Figure 1.39(b) we see an enlarged view of the nonzero portion of their product. By construction, $x_1(t)$ is zero outside the interval $0 \le t \le \Delta$. For Δ sufficiently small so that $x(t)$ is approximately constant over this interval,

$$x(t)\delta_\Delta(t) \approx x(0)\delta_\Delta(t).$$

Since $\delta(t)$ is the limit as $\Delta \to 0$ of $\delta_\Delta(t)$, it follows that

$$x(t)\delta(t) = x(0)\delta(t). \tag{1.76}$$

By the same argument, we have an analogous expression for an impulse concentrated at an arbitrary point, say, t_0. That is,

$$x(t)\delta(t - t_0) = x(t_0)\delta(t - t_0).$$

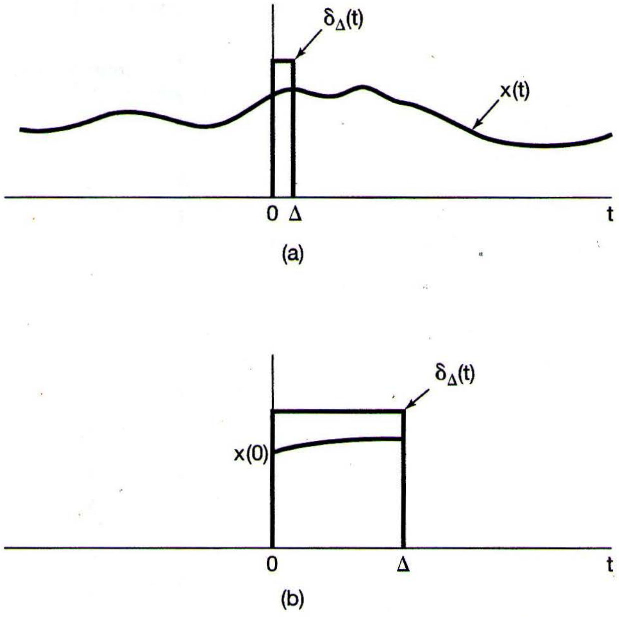


Figure 1.39 The product $x(t)\delta_\Delta(t)$: (a) graphs of both functions; (b) enlarged view of the nonzero portion of their product.

Although our discussion of the unit impulse in this section has been somewhat informal, it does provide us with some important intuition about this signal that will be useful throughout the book. As we have stated, the unit impulse should be viewed as an idealization. As we illustrate and discuss in more detail in Section 2.5, any real physical system has some inertia associated with it and thus does not respond instantaneously to inputs. Consequently, if a pulse of sufficiently short duration is applied to such a system, the system response will not be noticeably influenced by the pulse's duration or by the details of the shape of the pulse, for that matter. Instead, the primary characteristic of the pulse that will matter is the net, integrated effect of the pulse—i.e., its area. For systems that respond much more quickly than others, the pulse will have to be of much shorter duration before the details of the pulse shape or its duration no longer matter. Nevertheless, for any physical system, we can always find a pulse that is "short enough." The unit impulse then is an idealization of this concept—the pulse that is short enough for *any* system. As we will see in Chapter 2, the response of a system to this idealized pulse plays a crucial role in signal and system analysis, and in the process of developing and understanding this role, we will develop additional insight into the idealized signal.[3]

[3] The unit impulse and other related functions (which are often collectively referred to as *singularity functions*) have been thoroughly studied in the field of mathematics under the alternative names of *generalized functions* and the *theory of distributions.* For more detailed discussions of this subject, see *Distribution Theory and Transform Analysis,* by A. H. Zemanian (New York: McGraw-Hill Book Company, 1965), *Generalised Functions,* by R.F. Hoskins (New York: Halsted Press, 1979), or the more advanced text, *Fourier Analysis and Generalized Functions,* by M. J. Lighthill (New York: Cambridge University Press, 1958). Our discussion of singularity functions in Section 2.5 is closely related in spirit to the mathematical theory described in these texts and thus provides an informal introduction to concepts that underlie this topic in mathematics.

Example 1.7

Consider the discontinuous signal $x(t)$ depicted in Figure 1.40(a). Because of the relationship between the continuous-time unit impulse and unit step, we can readily calculate and graph the derivative of this signal. Specifically, the derivative of $x(t)$ is clearly 0, except at the discontinuities. In the case of the unit step, we have seen [eq. (1.72)] that differentiation gives rise to a unit impulse located at the point of discontinuity. Furthermore, by multiplying both sides of eq. (1.72) by any number k, we see that the derivative of a unit step with a discontinuity of size k gives rise to an impulse of area k at the point of discontinuity. This rule also holds for any other signal with a jump discontinuity, such as $x(t)$ in Figure 1.40(a). Consequently, we can sketch its derivative $\dot{x}(t)$, as in Figure 1.40(b), where an impulse is placed at each discontinuity of $x(t)$, with area equal to the size of the discontinuity. Note, for example, that the discontinuity in $x(t)$ at $t = 2$ has a value of -3, so that an impulse scaled by -3 is located at $t = 2$ in the signal $\dot{x}(t)$.

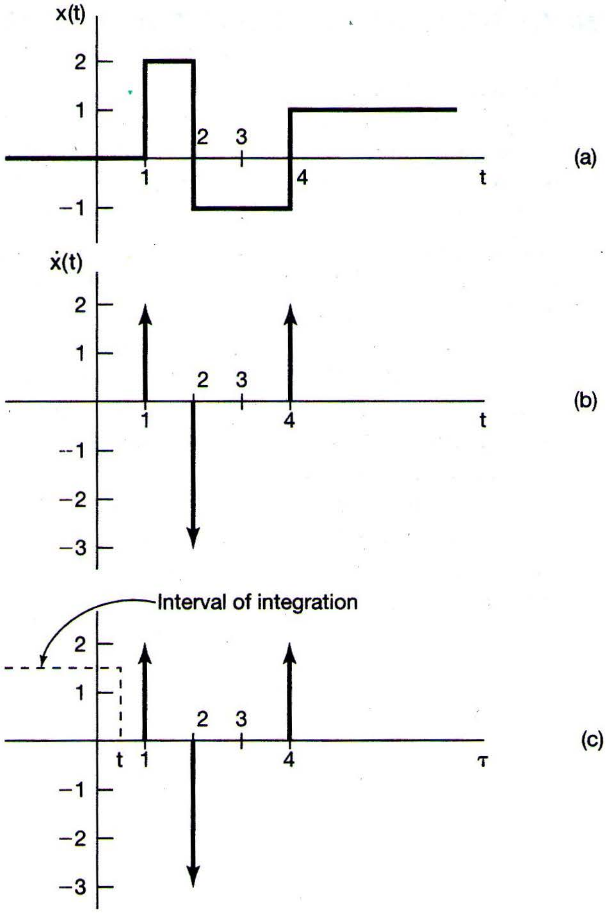


Figure 1.40 (a) The discontinuous signal $x(t)$ analyzed in Example 1.7; (b) its derivative $\dot{x}(t)$; (c) depiction of the recovery of $x(t)$ as the running integral of $\dot{x}(t)$, illustrated for a value of t between 0 and 1.

As a check of our result, we can verify that we can recover $x(t)$ from $\dot{x}(t)$. Specifically, since $x(t)$ and $\dot{x}(t)$ are both zero for $t \leq 0$, we need only check that for $t > 0$,

$$x(t) = \int_0^t \dot{x}(\tau)\, d\tau. \tag{1.77}$$

As illustrated in Figure 1.40(c), for $t < 1$, the integral on the right-hand side of eq. (1.77) is zero, since none of the impulses that constitute $\dot{x}(t)$ are within the interval of integration. For $1 < t < 2$, the first impulse (located at $t = 1$) is the only one within the integration interval, and thus the integral in eq. (1.77) equals 2, the area of this impulse. For $2 < t < 4$, the first two impulses are within the interval of integration, and the integral accumulates the sum of both of their areas, namely, $2 - 3 = -1$. Finally, for $t > 4$, all three impulses are within the integration interval, so that the integral equals the sum of all three areas—that is, $2 - 3 + 2 = +1$. The result is exactly the signal $x(t)$ depicted in Figure 1.40(a).

1.5 CONTINUOUS-TIME AND DISCRETE-TIME SYSTEMS

Physical systems in the broadest sense are an interconnection of components, devices, or subsystems. In contexts ranging from signal processing and communications to electromechanical motors, automotive vehicles, and chemical-processing plants, a *system* can be viewed as a process in which input signals are transformed by the system or cause the system to respond in some way, resulting in other signals as outputs. For example, a high-fidelity system takes a recorded audio signal and generates a reproduction of that signal. If the hi-fi system has tone controls, we can change the tonal quality of the reproduced signal. Similarly, the circuit in Figure 1.1 can be viewed as a system with input voltage $v_s(t)$ and output voltage $v_c(t)$, while the automobile in Figure 1.2 can be thought of as a system with input equal to the force $f(t)$ and output equal to the velocity $v(t)$ of the vehicle. An image-enhancement system transforms an input image into an output image that has some desired properties, such as improved contrast.

A *continuous-time system* is a system in which continuous-time input signals are applied and result in continuous-time output signals. Such a system will be represented pictorially as in Figure 1.41(a), where $x(t)$ is the input and $y(t)$ is the output. Alternatively, we will often represent the input-output relation of a continuous-time system by the notation

$$x(t) \rightarrow y(t). \tag{1.78}$$

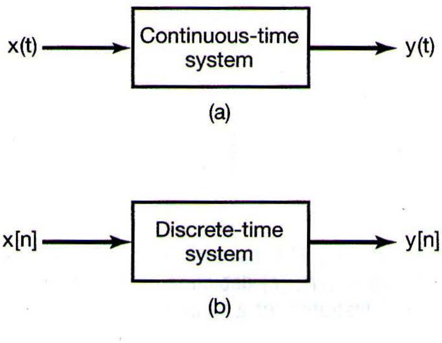


Figure 1.41 (a) Continuous-time system; (b) discrete-time system.

Similarly, a *discrete-time system*—that is, a system that transforms discrete-time inputs into discrete-time outputs—will be depicted as in Figure 1.41(b) and will sometimes be represented symbolically as

$$x[n] \rightarrow y[n]. \tag{1.79}$$

In most of this book, we will treat discrete-time systems and continuous-time systems separately but in parallel. In Chapter 7, we will bring continuous-time and discrete-time systems together through the concept of sampling, and we will develop some insights into the use of discrete-time systems to process continuous-time signals that have been sampled.

1.5.1 Simple Examples of Systems

One of the most important motivations for the development of general tools for analyzing and designing systems is that systems from many different applications have very similar mathematical descriptions. To illustrate this, we begin with a few simple examples.

Example 1.8

Consider the RC circuit depicted in Figure 1.1. If we regard $v_s(t)$ as the input signal and $v_c(t)$ as the output signal, then we can use simple circuit analysis to derive an equation describing the relationship between the input and output. Specifically, from Ohm's law, the current $i(t)$ through the resistor is proportional (with proportionality constant $1/R$) to the voltage drop across the resistor; i.e.,

$$i(t) = \frac{v_s(t) - v_c(t)}{R}. \tag{1.80}$$

Similarly, using the defining constitutive relation for a capacitor, we can relate $i(t)$ to the rate of change with time of the voltage across the capacitor:

$$i(t) = C\frac{dv_c(t)}{dt}. \tag{1.81}$$

Equating the right-hand sides of eqs. (1.80) and (1.81), we obtain a differential equation describing the relationship between the input $v_s(t)$ and the output $v_c(t)$:

$$\frac{dv_c(t)}{dt} + \frac{1}{RC}v_c(t) = \frac{1}{RC}v_s(t). \tag{1.82}$$

Example 1.9

Consider Figure 1.2, in which we regard the force $f(t)$ as the input and the velocity $v(t)$ as the output. If we let m denote the mass of the automobile and $m\rho v$ the resistance due to friction, then equating acceleration—i.e., the time derivative of velocity—with net force divided by mass, we obtain

$$\frac{dv(t)}{dt} = \frac{1}{m}\left[f(t) - \rho v(t)\right], \tag{1.83}$$

i.e.,

$$\frac{dv(t)}{dt} + \frac{\rho}{m}v(t) = \frac{1}{m}f(t). \tag{1.84}$$

Examining and comparing eqs. (1.82) and (1.84) in the above examples, we see that the input-output relationships captured in these two equations for these two very different physical systems are basically the same. In particular, they are both examples of first-order linear differential equations of the form

$$\frac{dy(t)}{dt} + ay(t) = bx(t), \tag{1.85}$$

where $x(t)$ is the input, $y(t)$ is the output, and a and b are constants. This is one very simple example of the fact that, by developing methods for analyzing general classes of systems such as that represented by eq. (1.85), we will be able to use them in a wide variety of applications.

Example 1.10

As a simple example of a discrete-time system, consider a simple model for the balance in a bank account from month to month. Specifically, let $y[n]$ denote the balance at the end of the nth month, and suppose that $y[n]$ evolves from month to month according to the equation

$$y[n] = 1.01y[n-1] + x[n], \tag{1.86}$$

or equivalently,

$$y[n] - 1.01y[n-1] = x[n], \tag{1.87}$$

where $x[n]$ represents the net deposit (i.e., deposits minus withdrawals) during the nth month and the term $1.01y[n-1]$ models the fact that we accrue 1% interest each month.

Example 1.11

As a second example, consider a simple digital simulation of the differential equation in eq. (1.84) in which we resolve time into discrete intervals of length Δ and approximate $dv(t)/dt$ at $t = n\Delta$ by the first backward difference, i.e.,

$$\frac{v(n\Delta) - v((n-1)\Delta)}{\Delta}.$$

In this case, if we let $v[n] = v(n\Delta)$ and $f[n] = f(n\Delta)$, we obtain the following discrete-time model relating the sampled signals $f[n]$ and $v[n]$:

$$v[n] - \frac{m}{(m+\rho\Delta)}v[n-1] = \frac{\Delta}{(m+\rho\Delta)}f[n]. \tag{1.88}$$

Comparing eqs. (1.87) and (1.88), we see that they are both examples of the same general first-order linear difference equation, namely,

$$y[n] + ay[n-1] = bx[n]. \tag{1.89}$$

As the preceding examples suggest, the mathematical descriptions of systems from a wide variety of applications frequently have a great deal in common, and it is this fact that provides considerable motivation for the development of broadly applicable tools for signal and system analysis. The key to doing this successfully is identifying classes of systems that have two important characteristics: (1) The systems in this class have properties and structures that we can exploit to gain insight into their behavior and to develop effective tools for their analysis; and (2) many systems of practical importance can be accurately modeled using systems in this class. It is on the first of these characteristics that most of this book focuses, as we develop tools for a particular class of systems referred to as linear, time-invariant systems. In the next section, we will introduce the properties that characterize this class, as well as a number of other very important basic system properties.

The second characteristic mentioned in the preceding paragraph is of obvious importance for any system analysis technique to be of value in practice. It is a well-established fact that a wide range of physical systems (including those in Examples 1.8–1.10) can be well modeled within the class of systems on which we focus in this book. However, a critical point is that *any* model used in describing or analyzing a physical system represents an idealization of that system, and thus, any resulting analysis is only as good as the model itself. For example, the simple linear model of a resistor in eq. (1.80) and that of a capacitor in eq. (1.81) are idealizations. However, these idealizations are quite accurate for real resistors and capacitors in many applications, and thus, analyses employing such idealizations provide useful results and conclusions, as long as the voltages and currents remain within the operating conditions under which these simple linear models are valid. Similarly, the use of a linear retarding force to represent frictional effects in eq. (1.83) is an approximation with a range of validity. Consequently, although we will not address this issue in the book, it is important to remember that an essential component of engineering practice in using the methods we develop here consists of identifying the range of validity of the assumptions that have gone into a model and ensuring that any analysis or design based on that model does not violate those assumptions.

1.5.2 Interconnections of Systems

An important idea that we will use throughout this book is the concept of the interconnection of systems. Many real systems are built as interconnections of several subsystems. One example is an audio system, which involves the interconnection of a radio receiver, compact disc player, or tape deck with an amplifier and one or more speakers. Another is a digitally controlled aircraft, which is an interconnection of the aircraft, described by its equations of motion and the aerodynamic forces affecting it; the sensors, which measure various aircraft variables such as accelerations, rotation rates, and heading; a digital autopilot, which responds to the measured variables and to command inputs from the pilot (e.g., the desired course, altitude, and speed); and the aircraft's actuators, which respond to inputs provided by the autopilot in order to use the aircraft control surfaces (rudder, tail, ailerons) to change the aerodynamic forces on the aircraft. By viewing such a system as an interconnection of its components, we can use our understanding of the component

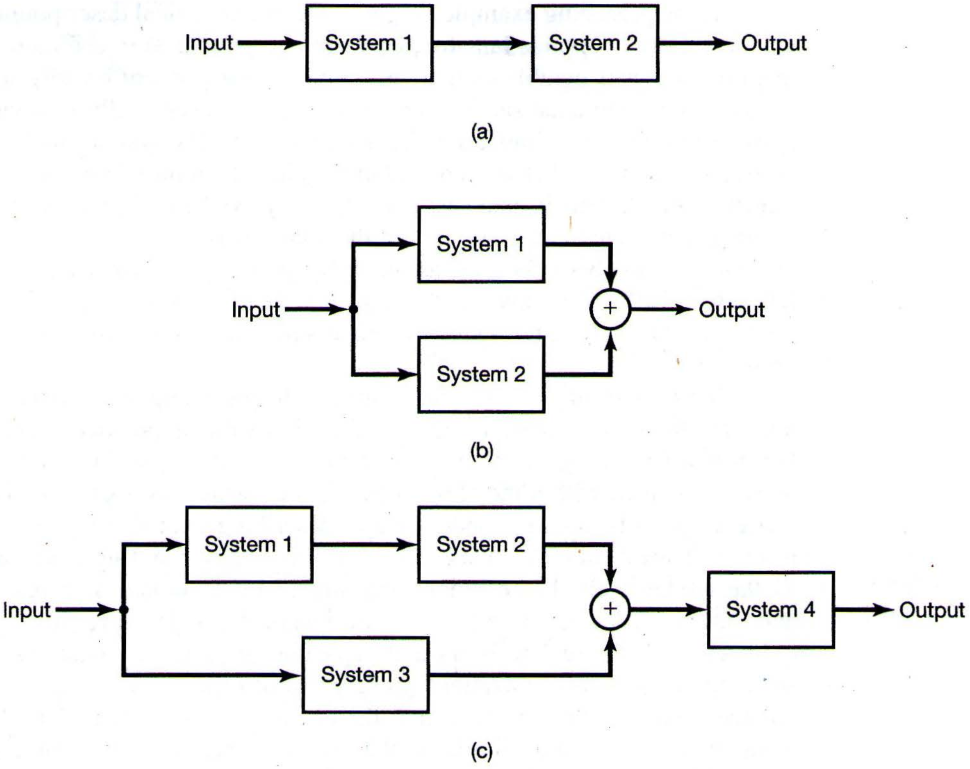


Figure 1.42 Interconnection of two systems: (a) series (cascade) interconnection; (b) parallel interconnection; (c) series-parallel interconnection.

systems and of how they are interconnected in order to analyze the operation and behavior of the overall system. In addition, by describing a system in terms of an interconnection of simpler subsystems, we may in fact be able to define useful ways in which to synthesize complex systems out of simpler, basic building blocks.

While one can construct a variety of system interconnections, there are several basic ones that are frequently encountered. A *series* or *cascade interconnection* of two systems is illustrated in Figure 1.42(a). Diagrams such as this are referred to as *block diagrams.* Here, the output of System 1 is the input to System 2, and the overall system transforms an input by processing it first by System 1 and then by System 2. An example of a series interconnection is a radio receiver followed by an amplifier. Similarly, one can define a series interconnection of three or more systems.

A *parallel interconnection* of two systems is illustrated in Figure 1.42(b). Here, the same input signal is applied to Systems 1 and 2. The symbol "⊕" in the figure denotes addition, so that the output of the parallel interconnection is the sum of the outputs of Systems 1 and 2. An example of a parallel interconnection is a simple audio system with several microphones feeding into a single amplifier and speaker system. In addition to the simple parallel interconnection in Figure 1.42(b), we can define parallel interconnections of more than two systems, and we can combine both cascade and parallel interconnections

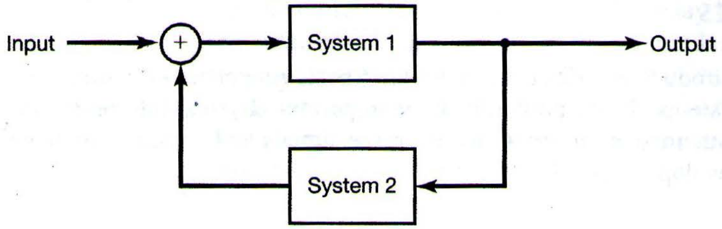


Figure 1.43 Feedback interconnection.

to obtain more complicated interconnections. An example of such an interconnection is given in Figure 1.42(c).[4]

Another important type of system interconnection is a *feedback interconnection,* an example of which is illustrated in Figure 1.43. Here, the output of System 1 is the input to System 2, while the output of System 2 is fed back and added to the external input to produce the actual input to System 1. Feedback systems arise in a wide variety of applications. For example, a cruise control system on an automobile senses the vehicle's velocity and adjusts the fuel flow in order to keep the speed at the desired level. Similarly, a digitally controlled aircraft is most naturally thought of as a feedback system in which differences between actual and desired speed, heading, or altitude are fed back through the autopilot in order to correct these discrepancies. Also, electrical circuits are often usefully viewed as containing feedback interconnections. As an example, consider the circuit depicted in Figure 1.44(a). As indicated in Figure 1.44(b), this system can be viewed as the feedback interconnection of the two circuit elements.

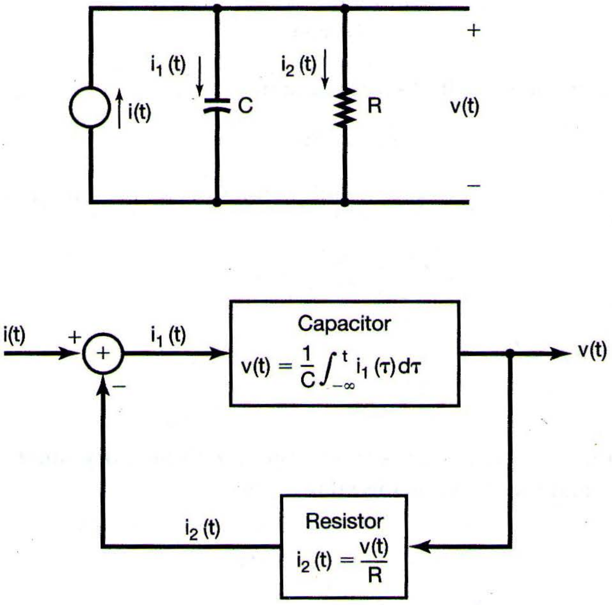


Figure 1.44 (a) Simple electrical circuit; (b) block diagram in which the circuit is depicted as the feedback interconnection of two circuit elements.

---

[4] On occasion, we will also use the symbol ⊗ in our pictorial representation of systems to denote the operation of multiplying two signals (see, for example, Figure 4.26).

1.6 BASIC SYSTEM PROPERTIES

In this section we introduce and discuss a number of basic properties of continuous-time and discrete-time systems. These properties have important physical interpretations and relatively simple mathematical descriptions using the signals and systems language that we have begun to develop.

1.6.1 Systems with and without Memory

A system is said to be *memoryless* if its output for each value of the independent variable at a given time is dependent on the input at only that same time. For example, the system specified by the relationship

$$y[n] = (2x[n] - x^2[n])^2 \tag{1.90}$$

is memoryless, as the value of $y[n]$ at any particular time n_0 depends only on the value of $x[n]$ at that time. Similarly, a resistor is a memoryless system; with the input $x(t)$ taken as the current and with the voltage taken as the output $y(t)$, the input-output relationship of a resistor is

$$y(t) = Rx(t), \tag{1.91}$$

where R is the resistance. One particularly simple memoryless system is the *identity system,* whose output is identical to its input. That is, the input-output relationship for the continuous-time identity system is

$$y(t) = x(t),$$

and the corresponding relationship in discrete time is

$$y[n] = x[n].$$

An example of a discrete-time system with memory is an *accumulator* or *summer*

$$y[n] = \sum_{k=-\infty}^{n} x[k], \tag{1.92}$$

and a second example is a *delay*

$$y[n] = x[n-1]. \tag{1.93}$$

A capacitor is an example of a continuous-time system with memory, since if the input is taken to be the current and the output is the voltage, then

$$y(t) = \frac{1}{C}\int_{-\infty}^{t} x(\tau)\,d\tau, \tag{1.94}$$

where C is the capacitance.

Roughly speaking, the concept of memory in a system corresponds to the presence of a mechanism in the system that retains or stores information about input values at times other than the current time. For example, the delay in eq. (1.93) must retain or store the preceding value of the input. Similarly, the accumulator in eq. (1.92) must "remember" or store information about past inputs. In particular, the accumulator computes the running sum of all inputs up to the current time, and thus, at each instant of time, the accumulator must add the current input value to the preceding value of the running sum. In other words, the relationship between the input and output of an accumulator can be described as

$$y[n] = \sum_{k=-\infty}^{n-1} x[k] + x[n], \tag{1.95}$$

or equivalently,

$$y[n] = y[n-1] + x[n]. \tag{1.96}$$

Represented in the latter way, to obtain the output at the current time n, the accumulator must remember the running sum of previous input values, which is exactly the preceding value of the accumulator output.

In many physical systems, memory is directly associated with the storage of energy. For example, the capacitor in eq. (1.94) stores energy by accumulating electrical charge, represented as the integral of the current. Thus, the simple *RC* circuit in Example 1.8 and Figure 1.1 has memory physically stored in the capacitor. Similarly, the automobile in Figure 1.2 has memory stored in its kinetic energy. In discrete-time systems implemented with computers or digital microprocessors, memory is typically directly associated with storage registers that retain values between clock pulses.

While the concept of memory in a system would typically suggest storing *past* input and output values, our formal definition also leads to our referring to a system as having memory if the current output is dependent on *future* values of the input and output. While systems having this dependence on future values might at first seem unnatural, they in fact form an important class of systems, as we discuss further in Section 1.6.3.

1.6.2 Invertibility and Inverse Systems

A system is said to be *invertible* if distinct inputs lead to distinct outputs. As illustrated in Figure 1.45(a) for the discrete-time case, if a system is invertible, then an *inverse system* exists that, when cascaded with the original system, yields an output $w[n]$ equal to the input $x[n]$ to the first system. Thus, the series interconnection in Figure 1.45(a) has an overall input-output relationship which is the same as that for the identity system.

An example of an invertible continuous-time system is

$$y(t) = 2x(t), \tag{1.97}$$

for which the inverse system is

$$w(t) = \frac{1}{2}y(t). \tag{1.98}$$

This example is illustrated in Figure 1.45(b). Another example of an invertible system is the accumulator of eq. (1.92). For this system, the difference between two successive

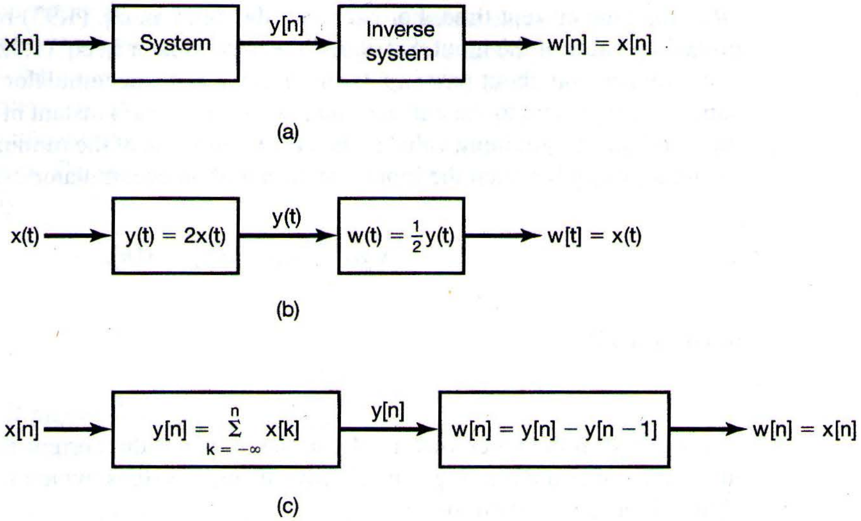


Figure 1.45 Concept of an inverse system for: (a) a general invertible system; (b) the invertible system described by eq. (1.97); (c) the invertible system defined in eq. (1.92).

values of the output is precisely the last input value. Therefore, in this case, the inverse system is

$$w[n] = y[n] - y[n-1], \tag{1.99}$$

as illustrated in Figure 1.45(c). Examples of noninvertible systems are

$$y[n] = 0, \tag{1.100}$$

that is, the system that produces the zero output sequence for any input sequence, and

$$y(t) = x^2(t), \tag{1.101}$$

in which case we cannot determine the sign of the input from knowledge of the output.

The concept of invertibility is important in many contexts. One example arises in systems for encoding used in a wide variety of communications applications. In such a system, a signal that we wish to transmit is first applied as the input to a system known as an encoder. There are many reasons for doing this, ranging from the desire to encrypt the original message for secure or private communication to the objective of providing some redundancy in the signal (for example, by adding what are known as parity bits) so that any errors that occur in transmission can be detected and, possibly, corrected. For *lossless* coding, the input to the encoder must be exactly recoverable from the output; i.e., the encoder must be invertible.

1.6.3 Causality

A system is *causal* if the output at any time depends on values of the input at only the present and past times. Such a system is often referred to as being *nonanticipative*, as the system output does not anticipate future values of the input. Consequently, if two inputs to a causal system are identical up to some point in time t_0 or n_0, the corresponding outputs must also be equal up to this same time. The *RC* circuit of Figure 1.1 is causal, since the capacitor voltage responds only to the present and past values of the source voltage. Similarly, the motion of an automobile is causal, since it does not anticipate future actions of the driver. The systems described in eqs. (1.92) – (1.94) are also causal, but the systems defined by

$$y[n] = x[n] - x[n+1] \tag{1.102}$$

and

$$y(t) = x(t+1) \tag{1.103}$$

are not. All memoryless systems are causal, since the output responds only to the current value of the input.

Although causal systems are of great importance, they do not by any means constitute the only systems that are of practical significance. For example, causality is not often an essential constraint in applications in which the independent variable is not time, such as in image processing. Furthermore, in processing data that have been recorded previously, as often happens with speech, geophysical, or meteorological signals, to name a few, we are by no means constrained to causal processing. As another example, in many applications, including historical stock market analysis and demographic studies, we may be interested in determining a slowly varying trend in data that also contain high-frequency fluctuations about that trend. In this case, a commonly used approach is to average data over an interval in order to smooth out the fluctuations and keep only the trend. An example of a noncausal averaging system is

$$y[n] = \frac{1}{2M+1} \sum_{k=-M}^{+M} x[n-k]. \tag{1.104}$$

Example 1.12

When checking the causality of a system, it is important to look carefully at the input-output relation. To illustrate some of the issues involved in doing this, we will check the causality of two particular systems.

The first system is defined by

$$y[n] = x[-n]. \tag{1.105}$$

Note that the output $y[n_0]$ at a positive time n_0 depends only on the value of the input signal $x[-n_0]$ at time $(-n_0)$, which is negative and therefore in the past of n_0. We may be tempted to conclude at this point that the given system is causal. However, we should always be careful to check the input-output relation for *all* times. In particular, for $n < 0$, e.g. $n = -4$, we see that $y[-4] = x[4]$, so that the output at this time depends on a future value of the input. Hence, the system is not causal.

It is also important to distinguish carefully the effects of the input from those of any other functions used in the definition of the system. For example, consider the system

$$y(t) = x(t)\cos(t+1). \tag{1.106}$$

In this system, the output at any time t equals the input at that same time multiplied by a number that varies with time. Specifically, we can rewrite eq. (1.106) as

$$y(t) = x(t)g(t),$$

where g(t) is a time-varying function, namely $g(t) = \cos(t + 1)$. Thus, only the current value of the input $x(t)$ influences the current value of the output $y(t)$, and we conclude that this system is causal (and, in fact, memoryless).

1.6.4 Stability

Stability is another important system property. Informally, a stable system is one in which small inputs lead to responses that do not diverge. For example, consider the pendulum in Figure 1.46(a), in which the input is the applied force $x(t)$ and the output is the angular deviation $y(t)$ from the vertical. In this case, gravity applies a restoring force that tends to return the pendulum to the vertical position, and frictional losses due to drag tend to slow it down. Consequently, if a small force $x(t)$ is applied, the resulting deflection from vertical will also be small. In contrast, for the inverted pendulum in Figure 1.46(b), the effect of gravity is to apply a force that tends to *increase* the deviation from vertical. Thus, a small applied force leads to a large vertical deflection causing the pendulum to topple over, despite any retarding forces due to friction.

The system in Figure 1.46(a) is an example of a stable system, while that in Figure 1.46(b) is unstable. Models for chain reactions or for population growth with unlimited food supplies and no predators are examples of unstable systems, since the system response grows without bound in response to small inputs. Another example of an unstable system is the model for a bank account balance in eq. (1.86), since if an initial deposit is made (i.e., $x[0]$ = a positive amount) and there are no subsequent withdrawals, then that deposit will grow each month without bound, because of the compounding effect of interest payments.

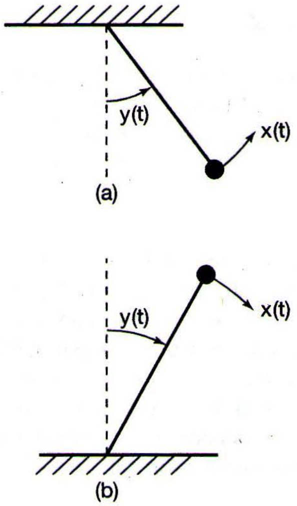


Figure 1.46 (a) A stable pendulum; (b) an unstable inverted pendulum.

There are also numerous examples of stable systems. Stability of physical systems generally results from the presence of mechanisms that dissipate energy. For example, assuming positive component values in the simple *RC* circuit of Example 1.8, the resistor dissipates energy and this circuit is a stable system. The system in Example 1.9 is also stable because of the dissipation of energy through friction.

The preceding examples provide us with an intuitive understanding of the concept of stability. More formally, if the input to a stable system is bounded (i.e., if its magnitude does not grow without bound), then the output must also be bounded and therefore cannot diverge. This is the definition of stability that we will use throughout this book. For example, consider applying a constant force $f(t) = F$ to the automobile in Figure 1.2, with the vehicle initially at rest. In this case the velocity of the car will increase, but not without bound, since the retarding frictional force also increases with velocity. In fact, the velocity will continue to increase until the frictional force exactly balances the applied force; so, from eq. (1.84), we see that this terminal velocity value V must satisfy

$$\frac{\rho}{m}V = \frac{1}{m}F, \tag{1.107}$$

i.e.,

$$V = \frac{F}{\rho}. \tag{1.108}$$

As another example, consider the discrete-time system defined by eq. (1.104), and suppose that the input $x[n]$ is bounded in magnitude by some number, say, B, for all values of n. Then the largest possible magnitude for $y[n]$ is also B, because $y[n]$ is the average of a finite set of values of the input. Therefore, $y[n]$ is bounded and the system is stable. On the other hand, consider the accumulator described by eq. (1.92). Unlike the system in eq. (1.104), this system sums *all* of the past values of the input rather than just a finite set of values, and the system is unstable, since the sum can grow continually even if $x[n]$ is bounded. For example, if the input to the accumulator is a unit step $u[n]$, the output will be

$$y[n] = \sum_{k=-\infty}^{n} u[k] = (n+1)u[n].$$

That is, $y[0] = 1$, $y[1] = 2$, $y[2] = 3$, and so on, and $y[n]$ grows without bound.

Example 1.13

If we suspect that a system is unstable, then a useful strategy to verify this is to look for a *specific* bounded input that leads to an unbounded output. Finding one such example enables us to conclude that the given system is unstable. If such an example does not exist or is difficult to find, we must check for stability by using a method that does not utilize specific examples of input signals. To illustrate this approach, let us check the stability of two systems,

$$S_1\colon y(t) = tx(t) \tag{1.109}$$

and

$$S_2\colon y(t) = e^{x(t)}. \tag{1.110}$$

In seeking a specific counterexample in order to disprove stability, we might try simple bounded inputs such as a constant or a unit step. For system S_1 in eq. (1.109), a constant input $x(t) = 1$ yields $y(t) = t$, which is unbounded, since no matter what finite constant we pick, $|y(t)|$ will exceed that constant for some t. We conclude that system S_1 is unstable.

For system S_2, which happens to be stable, we would be unable to find a bounded input that results in an unbounded output. So we proceed to verify that all bounded inputs result in bounded outputs. Specifically, let B be an arbitrary positive number, and let $x(t)$ be an arbitrary signal bounded by B; that is, we are making no assumption about $x(t)$, except that

$$|x(t)| < B, \tag{1.111}$$

or

$$-B < x(t) < B, \tag{1.112}$$

for all t. Using the definition of S_2 in eq. (1.110), we then see that if $x(t)$ satisfies eq. (1.111), then $y(t)$ must satisfy

$$e^{-B} < |y(t)| < e^{B}. \tag{1.113}$$

We conclude that if any input to S_2 is bounded by an arbitrary positive number B, the corresponding output is guaranteed to be bounded by e^B. Thus, S_2 is stable.

The system properties and concepts that we have introduced so far in this section are of great importance, and we will examine some of these in more detail later in the book. There remain, however, two additional properties—time invariance and linearity—that play a particularly central role in the subsequent chapters of the book, and in the remainder of this section we introduce and provide initial discussions of these two very important concepts.

1.6.5 Time Invariance

Conceptually, a system is time invariant if the behavior and characteristics of the system are fixed over time. For example, the *RC* circuit of Figure 1.1 is time invariant if the resistance and capacitance values R and C are constant over time: We would expect to get the same results from an experiment with this circuit today as we would if we ran the identical experiment tomorrow. On the other hand, if the values of R and C are changed or fluctuate over time, then we would expect the results of our experiment to depend on the time at which we run it. Similarly, if the frictional coefficient b and mass m of the automobile in Figure 1.2 are constant, we would expect the vehicle to respond identically independently of when we drive it. On the other hand, if we load the auto's trunk with heavy suitcases one day, thus increasing m, we would expect the car to behave differently than at other times when it is not so heavily loaded.

The property of time invariance can be described very simply in terms of the signals and systems language that we have introduced. Specifically, a system is time invariant if a time shift in the input signal results in an identical time shift in the output signal. That is, if $y[n]$ is the output of a discrete-time, time-invariant system when $x[n]$ is the input, then $y[n-n_0]$ is the output when $x[n-n_0]$ is applied. In continuous time with $y(t)$ the output corresponding to the input $x(t)$, a time-invariant system will have $y(t-t_0)$ as the output when $x(t-t_0)$ is the input.

To see how to determine whether a system is time invariant or not, and to gain some insight into this property, consider the following examples:

Example 1.14

Consider the continuous-time system defined by

$$y(t) = \sin[x(t)]. \tag{1.114}$$

To check that this system is time invariant, we must determine whether the time-invariance property holds for *any* input and *any* time shift t_0. Thus, let $x_1(t)$ be an arbitrary input to this system, and let

$$y_1(t) = \sin[x_1(t)] \tag{1.115}$$

be the corresponding output. Then consider a second input obtained by shifting $x_1(t)$ in time:

$$x_2(t) = x_1(t-t_0). \tag{1.116}$$

The output corresponding to this input is

$$y_2(t) = \sin[x_2(t)] = \sin[x_1(t-t_0)]. \tag{1.117}$$

Similarly, from eq. (1.115),

$$y_1(t-t_0) = \sin[x_1(t-t_0)]. \tag{1.118}$$

Comparing eqs. (1.117) and (1.118), we see that $y_2(t) = y_1(t-t_0)$, and therefore, this system is time invariant.

Example 1.15

As a second example, consider the discrete-time system

$$y[n] = nx[n]. \tag{1.119}$$

This is a time-varying system, a fact that can be verified using the same formal procedure as that used in the preceding example (see Problem 1.28). However, when a system is suspected of being time varying, an approach to showing this that is often very useful is to seek a counterexample—i.e., to use our intuition to find an input signal for which the condition of time invariance is violated. In particular, the system in this example represents a system with a time-varying gain. For example, if we know that the current input value is 1, we cannot determine the current output value without knowing the current time.

Consequently, consider the input signal $x_1[n] = \delta[n]$, which yields an output $y_1[n]$ that is identically 0 (since $n\delta[n] = 0$). However, the input $x_2[n] = \delta[n-1]$ yields the output $y_2[n] = n\delta[n-1] = \delta[n-1]$. Thus, while $x_2[n]$ is a shifted version of $x_1[n]$, $y_2[n]$ is *not* a shifted version of $y_1[n]$.

While the system in the preceding example has a time-varying gain and as a result is a time-varying system, the system in eq. (1.97) has a constant gain and, in fact, is time invariant. Other examples of time-invariant systems are given by eqs. (1.91)–(1.104). The following example illustrates a time-varying system.

Example 1.16

Consider the system

$$y(t) = x(2t). \tag{1.120}$$

This system represents a time scaling. That is, $y(t)$ is a time-compressed (by a factor of 2) version of $x(t)$. Intuitively, then, any time shift in the input will also be compressed by a factor of 2, and it is for this reason that the system is not time invariant. To demonstrate this by counterexample, consider the input $x_1(t)$ shown in Figure 1.47(a) and the resulting output $y_1(t)$ depicted in Figure 1.47(b). If we then shift the input by 2—i.e., consider $x_2(t) = x_1(t-2)$, as shown in Figure 1.47(c)—we obtain the resulting output

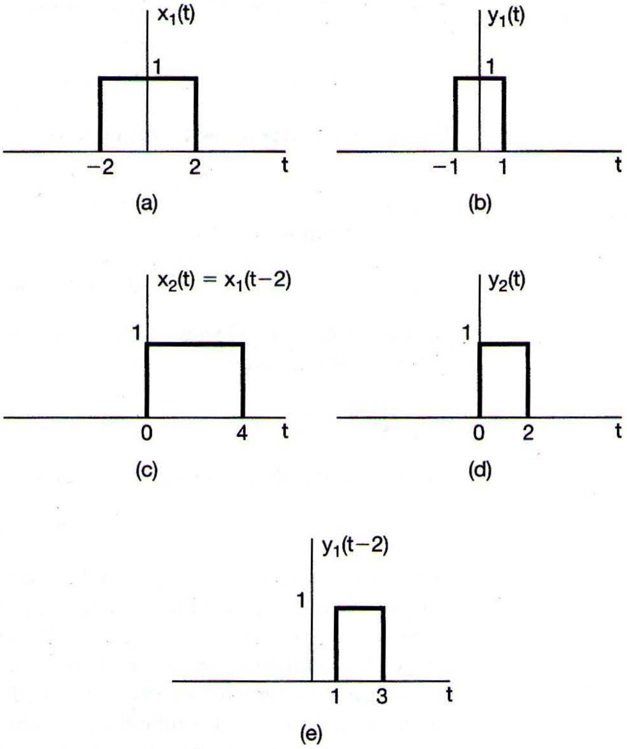


Figure 1.47 (a) The input $x_1(t)$ to the system in Example 1.16; (b) the output $y_1(t)$ corresponding to $x_1(t)$; (c) the shifted input $x_2(t) = x_1(t-2)$; (d) the output $y_2(t)$ corresponding to $x_2(t)$; (e) the shifted signal $y_1(t-2)$. Note that $y_2(t) \neq y_1(t-2)$, showing that the system is not time invariant.

$y_2(t) = x_2(2t)$ shown in Figure 1.47(d). Comparing Figures 1.47(d) and (e), we see that $y_2(t) \neq y_1(t-2)$, so that the system is not time invariant. (In fact, $y_2(t) = y_1(t-1)$, so that the output time shift is only half as big as it should be for time invariance, due to the time compression imparted by the system.)

1.6.6 Linearity

A *linear system,* in continuous time or discrete time, is a system that possesses the important property of superposition: If an input consists of the weighted sum of several signals, then the output is the superposition—that is, the weighted sum—of the responses of the system to each of those signals. More precisely, let $y_1(t)$ be the response of a continuous-time system to an input $x_1(t)$, and let $y_2(t)$ be the output corresponding to the input $x_2(t)$. Then the system is linear if:

1. The response to $x_1(t) + x_2(t)$ is $y_1(t) + y_2(t)$.
2. The response to $ax_1(t)$ is $ay_1(t)$, where a is any complex constant.

The first of these two properties is known as the *additivity* property; the second is known as the *scaling* or *homogeneity* property. Although we have written this description using continuous-time signals, the same definition holds in discrete time. The systems specified by eqs. (1.91)–(1.100), (1.102)–(1.104), and (1.119) are linear, while those defined by eqs. (1.101) and (1.114) are nonlinear. Note that a system can be linear without being time invariant, as in eq. (1.119), and it can be time invariant without being linear, as in eqs. (1.101) and (1.114).

The two properties defining a linear system can be combined into a single statement:

$$\text{continuous time: } ax_1(t) + bx_2(t) \rightarrow ay_1(t) + by_2(t), \tag{1.121}$$

$$\text{discrete time: } ax_1[n] + bx_2[n] \rightarrow ay_1[n] + by_2[n]. \tag{1.122}$$

Here, a and b are any complex constants. Furthermore, it is straightforward to show from the definition of linearity that if $x_k[n]$, $k = 1, 2, 3, \ldots$, are a set of inputs to a discrete-time linear system with corresponding outputs $y_k[n]$, $k = 1, 2, 3, \ldots$, then the response to a linear combination of these inputs given by

$$x[n] = \sum_k a_k x_k[n] = a_1 x_1[n] + a_2 x_2[n] + a_3 x_3[n] + \ldots \tag{1.123}$$

is

$$y[n] = \sum_k a_k y_k[n] = a_1 y_1[n] + a_2 y_2[n] + a_3 y_3[n] + \ldots. \tag{1.124}$$

This very important fact is known as the *superposition property*, which holds for linear systems in both continuous and discrete time.

A direct consequence of the superposition property is that, for linear systems, an input which is zero for all time results in an output which is zero for all time. For example, if $x[n] \rightarrow y[n]$, then the homogeneity property tells us that

$$0 = 0 \cdot x[n] \rightarrow 0 \cdot y[n] = 0. \tag{1.125}$$

In the following examples we illustrate how the linearity of a given system can be checked by directly applying the definition of linearity.

Example 1.17

Consider a system S whose input $x(t)$ and output $y(t)$ are related by

$$y(t) = tx(t)$$

To determine whether or not S is linear, we consider two arbitrary inputs $x_1(t)$ and $x_2(t)$.

$$\begin{aligned} x_1(t) &\to y_1(t) = tx_1(t) \\ x_2(t) &\to y_2(t) = tx_2(t) \end{aligned}$$

Let $x_3(t)$ be a linear combination of $x_1(t)$ and $x_2(t)$. That is,

$$x_3(t) = ax_1(t) + bx_2(t)$$

where a and b are arbitrary scalars. If $x_3(t)$ is the input to S, then the corresponding output may be expressed as

$$\begin{aligned} y_3(t) &= tx_3(t) \\ &= t(ax_1(t) + bx_2(t)) \\ &= atx_1(t) + btx_2(t) \\ &= ay_1(t) + by_2(t) \end{aligned}$$

We conclude that the system S is linear.

Example 1.18

Let us apply the linearity-checking procedure of the previous example to another system S whose input $x(t)$ and output $y(t)$ are related by

$$y(t) = x^2(t)$$

Defining $x_1(t)$, $x_2(t)$, and $x_3(t)$ as in the previous example, we have

$$\begin{aligned} x_1(t) &\to y_1(t) = x_1^2(t) \\ x_2(t) &\to y_2(t) = x_2^2(t) \end{aligned}$$

and

$$\begin{aligned} x_3(t) \to y_3(t) &= x_3^2(t) \\ &= (ax_1(t) + bx_2(t))^2 \\ &= a^2x_1^2(t) + b^2x_2^2(t) + 2abx_1(t)x_2(t) \\ &= a^2y_1(t) + b^2y_2(t) + 2abx_1(t)x_2(t) \end{aligned}$$

Clearly, we can specify $x_1(t)$, $x_2(t)$, a, and b such that $y_3(t)$ is not the same as $ay_1(t) + by_2(t)$. For example, if $x_1(t) = 1$, $x_2(t) = 0$, $a = 2$, and $b = 0$, then $y_3(t) = (2x_1(t))^2 = 4$, but $2y_1(t) = 2(x_1(t))^2 = 2$. We conclude that the system S is not linear.

Example 1.19

In checking the linearity of a system, it is important to remember that the system must satisfy both the additivity and homogeneity properties and that the signals, as well as any scaling constants, are allowed to be complex. To emphasize the importance of these points, consider the system specified by

$$y[n] = \mathcal{R}e\{x[n]\}. \tag{1.126}$$

As shown in Problem 1.29, this system is additive; however, it does not satisfy the homogeneity property, as we now demonstrate. Let

$$x_1[n] = r[n] + js[n] \tag{1.127}$$

be an arbitrary complex input with real and imaginary parts $r[n]$ and $s[n]$, respectively, so that the corresponding output is

$$y_1[n] = r[n]. \tag{1.128}$$

Now, consider scaling $x_1[n]$ by a complex number, for example, $a = j$; i.e., consider the input

$$\begin{aligned} x_2[n] &= jx_1[n] = j(r[n] + js[n]) \\ &= -s[n] + jr[n]. \end{aligned} \tag{1.129}$$

The output corresponding to $x_2[n]$ is

$$y_2[n] = \mathcal{R}e\{x_2[n]\} = -s[n], \tag{1.130}$$

which is not equal to the scaled version of $y_1[n]$,

$$ay_1[n] = jr[n]. \tag{1.131}$$

We conclude that the system violates the homogeneity property and hence is not linear.

Example 1.20

Consider the system

$$y[n] = 2x[n] + 3. \tag{1.132}$$

This system is not linear, as can be verified in several ways. For example, the system violates the additivity property: If $x_1[n] = 2$ and $x_2[n] = 3$, then

$$x_1[n] \rightarrow y_1[n] = 2x_1[n] + 3 = 7, \tag{1.133}$$

$$x_2[n] \rightarrow y_2[n] = 2x_2[n] + 3 = 9. \tag{1.134}$$

However, the response to $x_3[n] = x_1[n] + x_2[n]$ is

$$y_3[n] = 2[x_1[n] + x_2[n]] + 3 = 13, \tag{1.135}$$

which does not equal $y_1[n] + y_2[n] = 16$. Alternatively, since $y[n] = 3$ if $x[n] = 0$, we see that the system violates the "zero-in/zero-out" property of linear systems given in eq. (1.125).

It may seem surprising that the system in the above example is nonlinear, since eq. (1.132) is a linear equation. On the other hand, as depicted in Figure 1.48, the output of this system can be represented as the sum of the output of a linear system and another signal equal to the *zero-input response* of the system. For the system in eq. (1.132), the linear system is

$$x[n] \rightarrow 2x[n],$$

and the zero-input response is

$$y_0[n] = 3.$$

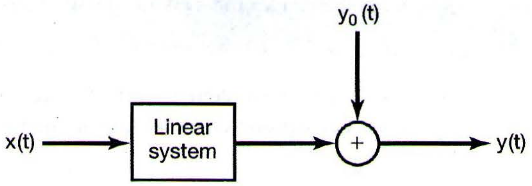


Figure 1.48 Structure of an incrementally linear system. Here, $y_0[n]$ is the zero-input response of the system.

There are, in fact, large classes of systems in both continuous and discrete time that can be represented as in Figure 1.48—i.e., for which the overall system output consists of the superposition of the response of a linear system with a zero-input response. As shown in Problem 1.47, such systems correspond to the class of *incrementally linear systems*—i.e., systems in continuous or discrete time that respond linearly to *changes* in the input. In other words, the *difference* between the responses to any two inputs to an incrementally linear system is a linear (i.e., additive and homogeneous) function of the *difference* between the two inputs. For example, if $x_1[n]$ and $x_2[n]$ are two inputs to the system specified by eq. (1.132), and if $y_1[n]$ and $y_2[n]$ are the corresponding outputs, then

$$y_1[n] - y_2[n] = 2x_1[n] + 3 - \{2x_2[n] + 3\} = 2\{x_1[n] - x_2[n]\}. \tag{1.136}$$

1.7 SUMMARY

In this chapter, we have developed a number of basic concepts related to continuous-time and discrete-time signals and systems. We have presented both an intuitive picture of what signals and systems are through several examples and a mathematical representation for signals and systems that we will use throughout the book. Specifically, we introduced graphical and mathematical representations of signals and used these representations in performing transformations of the independent variable. We also defined and examined several basic signals, both in continuous time and in discrete time. These included complex exponential signals, sinusoidal signals, and unit impulse and step functions. In addition, we investigated the concept of periodicity for continuous-time and discrete-time signals.

In developing some of the elementary ideas related to systems, we introduced block diagrams to facilitate our discussions concerning the interconnection of systems, and we defined a number of important properties of systems, including causality, stability, time invariance, and linearity.

The primary focus in most of this book will be on the class of linear, time-invariant (LTI) systems, both in continuous time and in discrete time. These systems play a particularly important role in system analysis and design, in part due to the fact that many systems encountered in nature can be successfully modeled as linear and time invariant. Furthermore, as we shall see in the following chapters, the properties of linearity and time invariance allow us to analyze in detail the behavior of LTI systems.

Chapter 1 Problems

Basic problems emphasize the mechanics of using concepts and methods in a manner similar to that illustrated in the examples that are solved in the text.

Advanced problems explore and elaborate upon the foundations and practical implications of the textual material.

The first section of problems belongs to the basic category, and the answers are provided in the back of the book. The next two sections contain problems belonging to the basic and advanced categories, respectively. A final section, **Mathematical Review**, provides practice problems on the fundamental ideas of complex arithmetic and algebra.

BASIC PROBLEMS WITH ANSWERS

1.1. Express each of the following complex numbers in Cartesian form ($x + jy$): $\frac{1}{2}e^{j\pi}$, $\frac{1}{2}e^{-j\pi}$, $e^{j\pi/2}$, $e^{-j\pi/2}$, $e^{j5\pi/2}$, $\sqrt{2}e^{j\pi/4}$, $\sqrt{2}e^{j9\pi/4}$, $\sqrt{2}e^{-j9\pi/4}$, $\sqrt{2}e^{-j\pi/4}$.

1.2. Express each of the following complex numbers in polar form ($re^{j\theta}$, with $-\pi < \theta \leq \pi$): $5, -2, -3j, \frac{1}{2} - j\frac{\sqrt{3}}{2}, 1+j, (1-j)^2, j(1-j), (1+j)/(1-j), (\sqrt{2} + j\sqrt{2})/(1 + j\sqrt{3})$.

1.3. Determine the values of P_∞ and E_∞ for each of the following signals:

(a) $x_1(t) = e^{-2t}u(t)$ **(b)** $x_2(t) = e^{j(2t+\pi/4)}$ **(c)** $x_3(t) = \cos(t)$

(d) $x_1[n] = (\frac{1}{2})^n u[n]$ **(e)** $x_2[n] = e^{j(\pi/2n+\pi/8)}$ **(f)** $x_3[n] = \cos(\frac{\pi}{4}n)$

1.4. Let $x[n]$ be a signal with $x[n] = 0$ for $n < -2$ and $n > 4$. For each signal given below, determine the values of n for which it is guaranteed to be zero.

(a) $x[n-3]$ **(b)** $x[n+4]$ **(c)** $x[-n]$

(d) $x[-n+2]$ **(e)** $x[-n-2]$

1.5. Let $x(t)$ be a signal with $x(t) = 0$ for $t < 3$. For each signal given below, determine the values of t for which it is guaranteed to be zero.

(a) $x(1-t)$ **(b)** $x(1-t) + x(2-t)$ **(c)** $x(1-t)x(2-t)$

(d) $x(3t)$ **(e)** $x(t/3)$

1.6. Determine whether or not each of the following signals is periodic:

(a) $x_1(t) = 2e^{j(t+\pi/4)}u(t)$ **(b)** $x_2[n] = u[n] + u[-n]$

(c) $x_3[n] = \sum_{k=-\infty}^{\infty}\{\delta[n-4k] - \delta[n-1-4k]\}$

1.7. For each signal given below, determine all the values of the independent variable at which the even part of the signal is guaranteed to be zero.

(a) $x_1[n] = u[n] - u[n-4]$ **(b)** $x_2(t) = \sin(\frac{1}{2}t)$

(c) $x_3[n] = (\frac{1}{2})^n u[n-3]$ **(d)** $x_4(t) = e^{-5t}u(t+2)$

1.8. Express the real part of each of the following signals in the form $Ae^{-at}\cos(\omega t + \phi)$, where A, a, ω, and ϕ are real numbers with $A > 0$ and $-\pi < \phi \leq \pi$:

(a) $x_1(t) = -2$ **(b)** $x_2(t) = \sqrt{2}e^{j\pi/4}\cos(3t + 2\pi)$

(c) $x_3(t) = e^{-t}\sin(3t + \pi)$ **(d)** $x_4(t) = je^{(-2+j100)t}$

1.9. Determine whether or not each of the following signals is periodic. If a signal is periodic, specify its fundamental period.

(a) $x_1(t) = je^{j10t}$ **(b)** $x_2(t) = e^{(-1+j)t}$ **(c)** $x_3[n] = e^{j7\pi n}$

(d) $x_4[n] = 3e^{j3\pi(n+1/2)/5}$ **(e)** $x_5[n] = 3e^{j3/5(n+1/2)}$

1.10. Determine the fundamental period of the signal $x(t) = 2\cos(10t+1) - \sin(4t-1)$.

1.11. Determine the fundamental period of the signal $x[n] = 1 + e^{j4\pi n/7} - e^{j2\pi n/5}$.

1.12. Consider the discrete-time signal

$$x[n] = 1 - \sum_{k=3}^{\infty} \delta[n-1-k].$$

Determine the values of the integers M and n_0 so that $x[n]$ may be expressed as

$$x[n] = u[Mn - n_0].$$

1.13. Consider the continuous-time signal

$$x(t) = \delta(t+2) - \delta(t-2).$$

Calculate the value of E_∞ for the signal

$$y(t) = \int_{-\infty}^{t} x(\tau)d\tau.$$

1.14. Consider a periodic signal

$$x(t) = \begin{cases} 1, & 0 \le t \le 1 \\ -2, & 1 < t < 2 \end{cases}$$

with period $T = 2$. The derivative of this signal is related to the "impulse train"

$$g(t) = \sum_{k=-\infty}^{\infty} \delta(t-2k)$$

with period $T = 2$. It can be shown that

$$\frac{dx(t)}{dt} = A_1 g(t-t_1) + A_2 g(t-t_2).$$

Determine the values of A_1, t_1, A_2, and t_2.

1.15. Consider a system S with input $x[n]$ and output $y[n]$. This system is obtained through a series interconnection of a system S_1 followed by a system S_2. The input-output relationships for S_1 and S_2 are

$$S_1: \quad y_1[n] = 2x_1[n] + 4x_1[n-1],$$

$$S_2: \quad y_2[n] = x_2[n-2] + \frac{1}{2}x_2[n-3],$$

where $x_1[n]$ and $x_2[n]$ denote input signals.

(a) Determine the input-output relationship for system S.

(b) Does the input-output relationship of system S change if the order in which S_1 and S_2 are connected in series is reversed (i.e., if S_2 follows S_1)?

1.16. Consider a discrete-time system with input $x[n]$ and output $y[n]$. The input-output relationship for this system is

$$y[n] = x[n]x[n-2].$$

(a) Is the system memoryless?
(b) Determine the output of the system when the input is $A\delta[n]$, where A is any real or complex number.
(c) Is the system invertible?

1.17. Consider a continuous-time system with input $x(t)$ and output $y(t)$ related by

$$y(t) = x(\sin(t)).$$

(a) Is this system causal?
(b) Is this system linear?

1.18. Consider a discrete-time system with input $x[n]$ and output $y[n]$ related by

$$y[n] = \sum_{k=n-n_0}^{n+n_0} x[k],$$

where n_0 is a finite positive integer.
(a) Is this system linear?
(a) Is this system time-invariant?
(c) If $x[n]$ is known to be bounded by a finite integer B (i.e., $|x[n]| < B$ for all n), it can be shown that $y[n]$ is bounded by a finite number C. We conclude that the given system is stable. Express C in terms of B and n_0.

1.19. For each of the following input-output relationships, determine whether the corresponding system is linear, time invariant or both.
(a) $y(t) = t^2x(t-1)$ **(b)** $y[n] = x^2[n-2]$
(c) $y[n] = x[n+1] - x[n-1]$ **(d)** $y[n] = \mathcal{O}d\{x(t)\}$

1.20. A continuous-time linear system S with input $x(t)$ and output $y(t)$ yields the following input-output pairs:

$$x(t) = e^{j2t} \xrightarrow{S} y(t) = e^{j3t},$$
$$x(t) = e^{-j2t} \xrightarrow{S} y(t) = e^{-j3t}.$$

(a) If $x_1(t) = \cos(2t)$, determine the corresponding output $y_1(t)$ for system S.
(b) If $x_2(t) = \cos(2(t - \frac{1}{2}))$, determine the corresponding output $y_2(t)$ for system S.

BASIC PROBLEMS

1.21. A continuous-time signal $x(t)$ is shown in Figure P1.21. Sketch and label carefully each of the following signals:
(a) $x(t-1)$ **(b)** $x(2-t)$ **(c)** $x(2t+1)$
(d) $x(4-\frac{t}{2})$ **(e)** $[x(t)+x(-t)]u(t)$ **(f)** $x(t)[\delta(t+\frac{3}{2}) - \delta(t-\frac{3}{2})]$

1.22. A discrete-time signal is shown in Figure P1.22. Sketch and label carefully each of the following signals:
(a) $x[n-4]$ **(b)** $x[3-n]$ **(c)** $x[3n]$
(d) $x[3n+1]$ **(e)** $x[n]u[3-n]$ **(f)** $x[n-2]\delta[n-2]$
(g) $\frac{1}{2}x[n] + \frac{1}{2}(-1)^n x[n]$ **(h)** $x[(n-1)^2]$

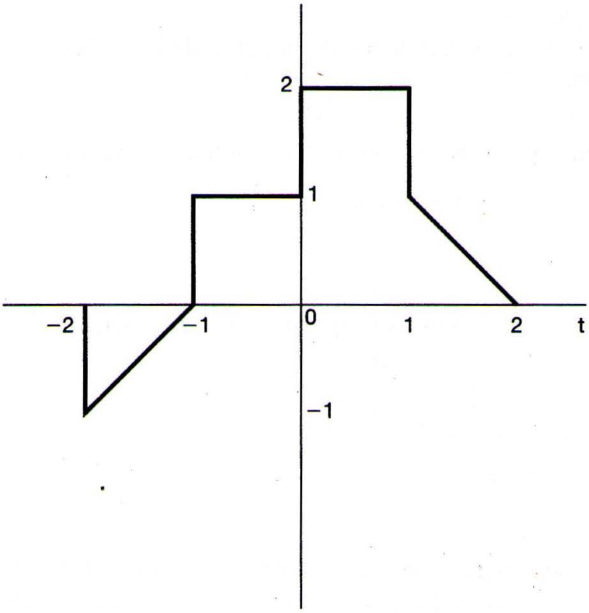


Figure P1.21

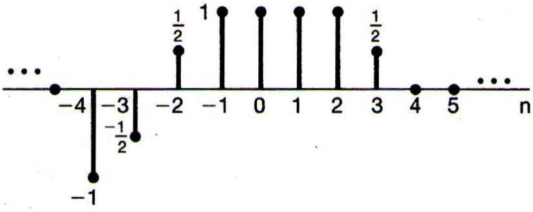


Figure P1.22

1.23. Determine and sketch the even and odd parts of the signals depicted in Figure P1.23. Label your sketches carefully.

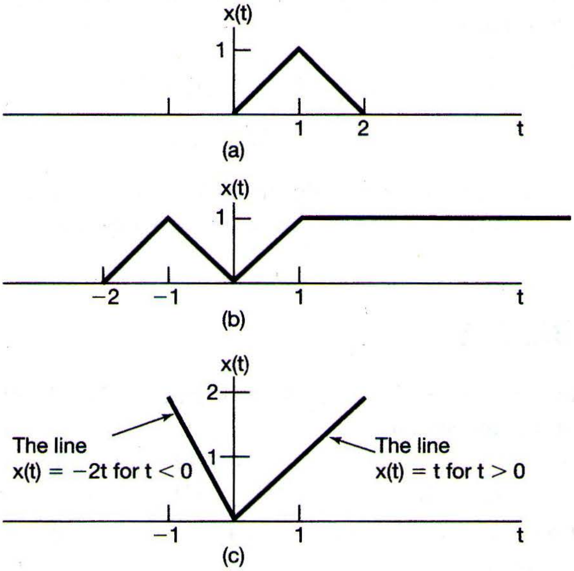
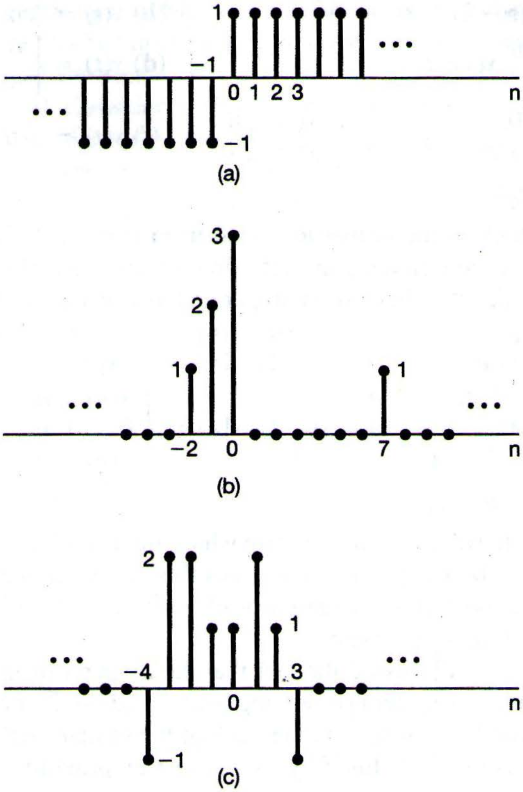


Figure P1.23

1.24. Determine and sketch the even and odd parts of the signals depicted in Figure P1.24. Label your sketches carefully.



(a)

(b)

(c)

Figure P1.24

1.25. Determine whether or not each of the following continuous-time signals is periodic. If the signal is periodic, determine its fundamental period.

(a) $x(t) = 3\cos(4t + \frac{\pi}{3})$ **(b)** $x(t) = e^{j(\pi t - 1)}$

(c) $x(t) = [\cos(2t - \frac{\pi}{3})]^2$ **(d)** $x(t) = \mathcal{E}v\{\cos(4\pi t)u(t)\}$

(e) $x(t) = \mathcal{E}v\{\sin(4\pi t)u(t)\}$ **(f)** $x(t) = \sum_{n=-\infty}^{\infty} e^{-(2t-n)}u(2t-n)$

1.26. Determine whether or not each of the following discrete-time signals is periodic. If the signal is periodic, determine its fundamental period.

(a) $x[n] = \sin(\frac{6\pi}{7}n + 1)$ **(b)** $x[n] = \cos(\frac{n}{8} - \pi)$ **(c)** $x[n] = \cos(\frac{\pi}{8}n^2)$

(d) $x[n] = \cos(\frac{\pi}{2}n)\cos(\frac{\pi}{4}n)$ **(e)** $x[n] = 2\cos(\frac{\pi}{4}n) + \sin(\frac{\pi}{8}n) - 2\cos(\frac{\pi}{2}n + \frac{\pi}{6})$

1.27. In this chapter, we introduced a number of general properties of systems. In particular, a system may or may not be

(1) Memoryless
(2) Time invariant
(3) Linear
(4) Causal
(5) Stable

Determine which of these properties hold and which do not hold for each of the following continuous-time systems. Justify your answers. In each example, $y(t)$ denotes the system output and $x(t)$ is the system input.

(a) $y(t) = x(t-2) + x(2-t)$ **(b)** $y(t) = [\cos(3t)]x(t)$

(c) $y(t) = \int_{-\infty}^{2t} x(\tau)d\tau$ **(d)** $y(t) = \begin{cases} 0, & t < 0 \\ x(t) + x(t-2), & t \geq 0 \end{cases}$

(e) $y(t) = \begin{cases} 0, & x(t) < 0 \\ x(t) + x(t-2), & x(t) \geq 0 \end{cases}$ **(f)** $y(t) = x(t/3)$

(g) $y(t) = \frac{dx(t)}{dt}$

1.28. Determine which of the properties listed in Problem 1.27 hold and which do not hold for each of the following discrete-time systems. Justify your answers. In each example, $y[n]$ denotes the system output and $x[n]$ is the system input.

(a) $y[n] = x[-n]$ **(b)** $y[n] = x[n-2] - 2x[n-8]$

(c) $y[n] = nx[n]$ **(d)** $y[n] = \mathcal{E}v\{x[n-1]\}$

(e) $y[n] = \begin{cases} x[n], & n \geq 1 \\ 0, & n = 0 \\ x[n+1], & n \leq -1 \end{cases}$ **(f)** $y[n] = \begin{cases} x[n], & n \geq 1 \\ 0, & n = 0 \\ x[n], & n \leq -1 \end{cases}$

(g) $y[n] = x[4n+1]$

1.29. **(a)** Show that the discrete-time system whose input $x[n]$ and output $y[n]$ are related by $y[n] = \mathcal{R}e\{x[n]\}$ is additive. Does this system remain additive if its input-output relationship is changed to $y[n] = \mathcal{R}e\{e^{j\pi n/4}x[n]\}$? (Do not assume that $x[n]$ is real in this problem.)

(b) In the text, we discussed the fact that the property of linearity for a system is equivalent to the system possessing both the additivity property and homogeneity property. Determine whether each of the systems defined below is additive and/or homogeneous. Justify your answers by providing a proof for each property if it holds or a counterexample if it does not.

(i) $y(t) = \frac{1}{x(t)}\left[\frac{dx(t)}{dt}\right]^2$ (ii) $y[n] = \begin{cases} \frac{x[n]x[n-2]}{x[n-1]}, & x[n-1] \neq 0 \\ 0, & x[n-1] = 0 \end{cases}$

1.30. Determine if each of the following systems is invertible. If it is, construct the inverse system. If it is not, find two input signals to the system that have the same output.

(a) $y(t) = x(t-4)$ **(b)** $y(t) = \cos[x(t)]$

(c) $y[n] = nx[n]$ **(d)** $y(t) = \int_{-\infty}^{t} x(\tau)d\tau$

(e) $y[n] = \begin{cases} x[n-1], & n \geq 1 \\ 0, & n = 0 \\ x[n], & n \leq -1 \end{cases}$ **(f)** $y[n] = x[n]x[n-1]$

(g) $y[n] = x[1-n]$ **(h)** $y(t) = \int_{-\infty}^{t} e^{-(t-\tau)}x(\tau)d\tau$

(i) $y[n] = \sum_{k=-\infty}^{n} (\frac{1}{2})^{n-k}x[k]$ **(j)** $y(t) = \frac{dx(t)}{dt}$

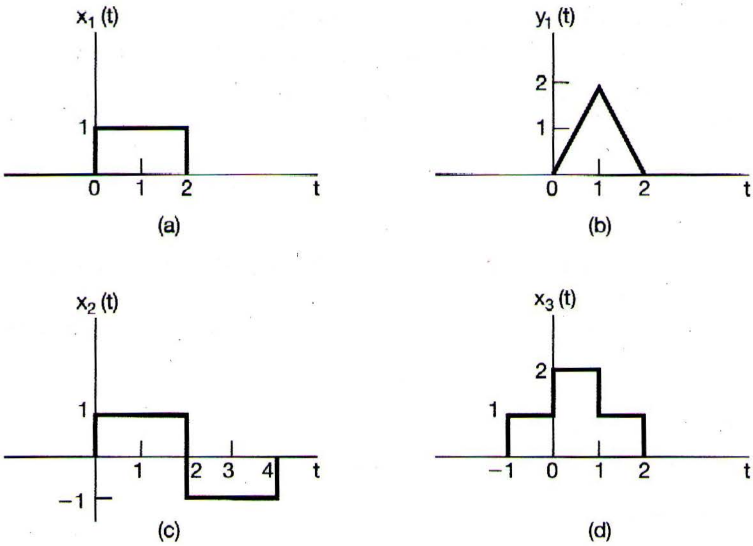
(k) $y[n] = \begin{cases} x[n+1], & n \geq 0 \\ x[n], & n \leq -1 \end{cases}$ **(l)** $y(t) = x(2t)$

(m) $y[n] = x[2n]$ **(n)** $y[n] = \begin{cases} x[n/2], & n \text{ even} \\ 0, & n \text{ odd} \end{cases}$

1.31. In this problem, we illustrate one of the most important consequences of the properties of linearity and time invariance. Specifically, once we know the response of a linear system or a linear time-invariant (LTI) system to a single input or the responses to several inputs, we can directly compute the responses to many other input signals. Much of the remainder of this book deals with a thorough exploitation of this fact in order to develop results and techniques for analyzing and synthesizing LTI systems.

(a) Consider an LTI system whose response to the signal $x_1(t)$ in Figure P1.31(a) is the signal $y_1(t)$ illustrated in Figure P1.31(b). Determine and sketch carefully the response of the system to the input $x_2(t)$ depicted in Figure P1.31(c).

(b) Determine and sketch the response of the system considered in part (a) to the input $x_3(t)$ shown in Figure P1.31(d).



(a) (b) (c) (d)

Figure P1.31

ADVANCED PROBLEMS

1.32. Let $x(t)$ be a continuous-time signal, and let

$$y_1(t) = x(2t) \text{ and } y_2(t) = x(t/2).$$

The signal $y_1(t)$ represents a speeded up version of $x(t)$ in the sense that the duration of the signal is cut in half. Similarly, $y_2(t)$ represents a slowed down version of $x(t)$ in the sense that the duration of the signal is doubled. Consider the following statements:

(1) If $x(t)$ is periodic, then $y_1(t)$ is periodic.
(2) If $y_1(t)$ is periodic, then $x(t)$ is periodic.
(3) If $x(t)$ is periodic, then $y_2(t)$ is periodic.
(4) If $y_2(t)$ is periodic, then $x(t)$ is periodic.

For each of these statements, determine whether it is true, and if so, determine the relationship between the fundamental periods of the two signals considered in the statement. If the statement is not true, produce a counterexample to it.

1.33. Let $x[n]$ be a discrete-time signal, and let

$$y_1[n] = x[2n] \text{ and } y_2[n] = \begin{cases} x[n/2], & n \text{ even} \\ 0, & n \text{ odd} \end{cases}.$$

The signals $y_1[n]$ and $y_2[n]$ respectively represent in some sense the speeded up and slowed down versions of $x[n]$. However, it should be noted that the discrete-time notions of speeded up and slowed down have subtle differences with respect to their continuous-time counterparts. Consider the following statements:

(1) If $x[n]$ is periodic, then $y_1[n]$ is periodic.
(2) If $y_1[n]$ is periodic, then $x[n]$ is periodic.
(3) If $x[n]$ is periodic, then $y_2[n]$ is periodic.
(4) If $y_2[n]$ is periodic, then $x[n]$ is periodic.

For each of these statements, determine whether it is true, and if so, determine the relationship between the fundamental periods of the two signals considered in the statement. If the statement is not true, produce a counterexample to it.

1.34. In this problem, we explore several of the properties of even and odd signals.

(a) Show that if $x[n]$ is an odd signal, then

$$\sum_{n=-\infty}^{+\infty} x[n] = 0.$$

(b) Show that if $x_1[n]$ is an odd signal and $x_2[n]$ is an even signal, then $x_1[n]x_2[n]$ is an odd signal.

(c) Let $x[n]$ be an arbitrary signal with even and odd parts denoted by

$$x_e[n] = \mathcal{E}v\{x[n]\}$$

and

$$x_o[n] = \mathcal{O}d\{x[n]\}.$$

Show that

$$\sum_{n=-\infty}^{+\infty} x^2[n] = \sum_{n=-\infty}^{+\infty} x_e^2[n] + \sum_{n=-\infty}^{+\infty} x_o^2[n].$$

(d) Although parts (a)–(c) have been stated in terms of discrete-time signals, the analogous properties are also valid in continuous time. To demonstrate this, show that

$$\int_{-\infty}^{+\infty} x^2(t)dt = \int_{-\infty}^{+\infty} x_e^2(t)dt + \int_{-\infty}^{+\infty} x_o^2(t)dt,$$

where $x_e(t)$ and $x_o(t)$ are, respectively, the even and odd parts of $x(t)$.

1.35. Consider the periodic discrete-time exponential time signal

$$x[n] = e^{jm(2\pi/N)n}.$$

Show that the fundamental period of this signal is

$$N_0 = N/\gcd(m, N),$$

where $\gcd(m, N)$ is the *greatest common divisor* of m and N—that is, the largest integer that divides both m and N an integral number of times. For example,

$$\gcd(2, 3) = 1, \gcd(2, 4) = 2, \gcd(8, 12) = 4.$$

Note that $N_0 = N$ if m and N have no factors in common.

1.36. Let $x(t)$ be the continuous-time complex exponential signal

$$x(t) = e^{j\omega_0 t}$$

with fundamental frequency ω_0 and fundamental period $T_0 = 2\pi/\omega_0$. Consider the discrete-time signal obtained by taking equally spaced samples of $x(t)$—that is,

$$x[n] = x(nT) = e^{j\omega_0 nT}.$$

(a) Show that $x[n]$ is periodic if and only if T/T_0 is a rational number—that is, if and only if some multiple of the sampling interval *exactly equals* a multiple of the period of $x(t)$.

(b) Suppose that $x[n]$ is periodic—that is, that

$$\frac{T}{T_0} = \frac{p}{q}, \tag{P1.36–1}$$

where p and q are integers. What are the fundamental period and fundamental frequency of $x[n]$? Express the fundamental frequency as a fraction of $\omega_0 T$.

(c) Again assuming that T/T_0 satisfies eq. (P1.36–1), determine precisely how many periods of $x(t)$ are needed to obtain the samples that form a single period of $x[n]$.

1.37. An important concept in many communications applications is the *correlation* between two signals. In the problems at the end of Chapter 2, we will have more to say about this topic and will provide some indication of how it is used in practice. For now, we content ourselves with a brief introduction to correlation functions and some of their properties.

Let $x(t)$ and $y(t)$ be two signals; then the *correlation function* is defined as

$$\phi_{xy}(t) = \int_{-\infty}^{\infty} x(t+\tau)y(\tau)d\tau.$$

The function $\phi_{xx}(t)$ is usually referred to as the *autocorrelation function* of the signal $x(t)$, while $\phi_{xy}(t)$ is often called a *cross-correlation function*.

(a) What is the relationship between $\phi_{xy}(t)$ and $\phi_{yx}(t)$?

(b) Compute the odd part of $\phi_{xx}(t)$.

(c) Suppose that $y(t) = x(t+T)$. Express $\phi_{xy}(t)$ and $\phi_{yy}(t)$ in terms of $\phi_{xx}(t)$.

1.38. In this problem, we examine a few of the properties of the unit impulse function.

(a) Show that

$$\delta(2t) = \frac{1}{2}\delta(t).$$

Hint: Examine $\delta_\Delta(t)$. (See Figure 1.34.)

(b) In Section 1.4, we defined the continuous-time unit impulse as the limit of the signal $\delta_\Delta(t)$. More precisely, we defined several of the *properties* of $\delta(t)$ by examining the corresponding properties of $\delta_\Delta(t)$. For example, since the signal

$$u_\Delta(t) = \int_{-\infty}^{t} \delta_\Delta(\tau)d\tau$$

converges to the unit step

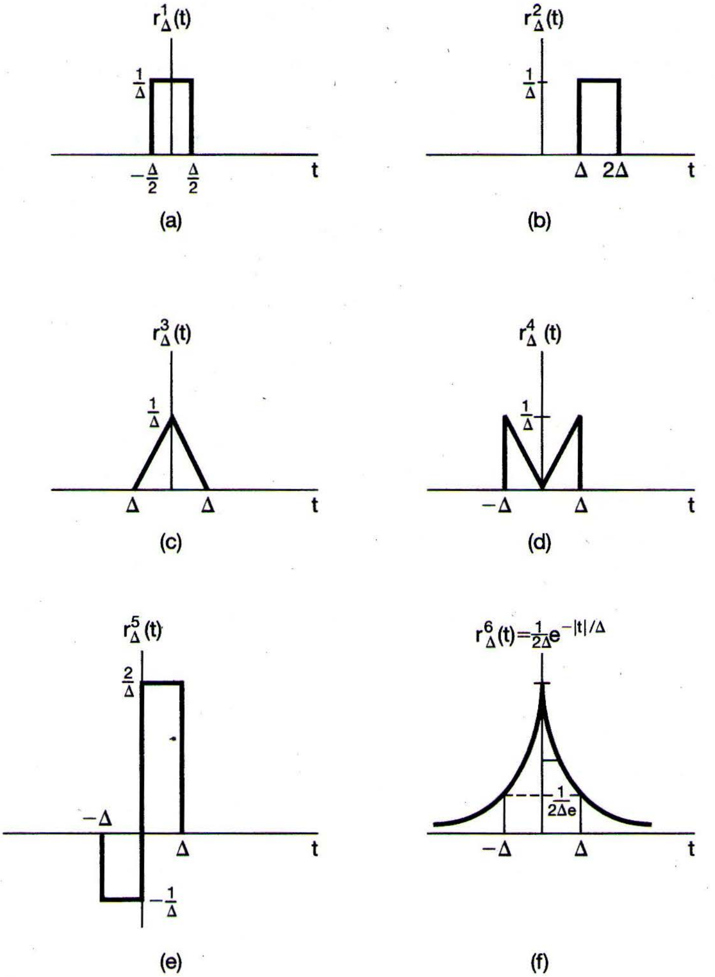
$$u(t) = \lim_{\Delta \to 0} u_\Delta(t), \qquad \text{(P1.38–1)}$$

we could interpret $\delta(t)$ through the equation

$$u(t) = \int_{-\infty}^{t} \delta(\tau)d\tau$$

or by viewing $\delta(t)$ as the formal derivative of $u(t)$.

This type of discussion is important, as we are in effect trying to define $\delta(t)$ through its properties rather than by specifying its value for each t, which is not possible. In Chapter 2, we provide a very simple characterization of the behavior of the unit impulse that is extremely useful in the study of linear time-invariant systems. For the present, however, we concentrate on demonstrating that the important concept in using the unit impulse is to understand *how* it behaves. To do this, consider the six signals depicted in Figure P1.38. Show



$r^1_\Delta(t)$
$\frac{1}{\Delta}$
$-\frac{\Delta}{2}$
$\frac{\Delta}{2}$
t
(a)
$r^2_\Delta(t)$
$\frac{1}{\Delta}$
Δ
2Δ
t
(b)
$r^3_\Delta(t)$
$\frac{1}{\Delta}$
Δ
Δ
t
(c)
$r^4_\Delta(t)$
$\frac{1}{\Delta}$
−Δ
Δ
t
(d)
$r^5_\Delta(t)$
$\frac{2}{\Delta}$
−Δ
Δ
$-\frac{1}{\Delta}$
t
(e)
$r^6_\Delta(t) = \frac{1}{2\Delta}e^{-|t|/\Delta}$
$\frac{1}{2\Delta e}$
−Δ
Δ
t
(f)

Figure P1.38

that each "behaves like an impulse" as $\Delta \rightarrow 0$ in that, if we let

$$u_\Delta^i(t) = \int_{-\infty}^{t} r_\Delta^i(\tau)d\tau,$$

then

$$\lim_{\Delta \to 0} u_\Delta^i(t) = u(t).$$

In each case, sketch and label carefully the signal $u_\Delta^i(t)$. Note that

$$r_\Delta^2(0) = r_\Delta^4(0) = 0 \text{ for all } \Delta.$$

Therefore, it is not enough to define or to think of $\delta(t)$ as being zero for $t \neq 0$ and infinite for $t = 0$. Rather, it is properties such as eq. (P1.38–1) that define the impulse. In Section 2.5 we will define a whole class of signals known as *singularity functions*, which are related to the unit impulse and which are also defined in terms of their properties rather than their values.

1.39. The role played by $u(t)$, $\delta(t)$, and other singularity functions in the study of linear time-invariant systems is that of an *idealization* of a physical phenomenon, and, as we will see, the use of these idealizations allow us to obtain an exceedingly important and very simple representation of such systems. In using singularity functions, we need, however, to be careful. In particular, we must remember that they are idealizations, and thus, whenever we perform a calculation using them, we are implicitly assuming that this calculation represents an accurate description of the behavior of the signals that they are intended to idealize. To illustrate, consider the equation

$$x(t)\delta(t) = x(0)\delta(t). \tag{P1.39–1}$$

This equation is based on the observation that

$$x(t)\delta_\Delta(t) \approx x(0)\delta_\Delta(t). \tag{P1.39–2}$$

Taking the limit of this relationship then yields the idealized one given by eq. (P1.39–1). However, a more careful examination of our derivation of eq. (P1.39–2) shows that that equation really makes sense only if $x(t)$ is continuous at $t = 0$. If it is not, then we will not have $x(t) \approx x(0)$ for t small.

To make this point clearer, consider the unit step signal $u(t)$. Recall from eq. (1.70) that $u(t) = 0$ for $t < 0$ and $u(t) = 1$ for $t > 0$, but that its value at $t = 0$ is not defined. [Note, for example, that $u_\Delta(0) = 0$ for all Δ, while $u_\Delta^1(0) = \frac{1}{2}$ (from Problem 1.38(b)).] The fact that $u(0)$ is not defined is not particularly bothersome, as long as the calculations we perform using $u(t)$ do not rely on a specific choice for $u(0)$. For example, if $f(t)$ is a signal that is continuous at $t = 0$, then the value of

$$\int_{-\infty}^{+\infty} f(\sigma)u(\sigma)d\sigma$$

does not depend upon a choice for $u(0)$. On the other hand, the fact that $u(0)$ is undefined is significant in that it means that certain calculations involving singularity functions are undefined. Consider trying to define a value for the product $u(t)\delta(t)$.

To see that this *cannot* be defined, show that

$$\lim_{\Delta \to 0} [u_\Delta(t)\delta(t)] = 0,$$

but

$$\lim_{\Delta \to 0} [u_\Delta(t)\delta_\Delta(t)] = \frac{1}{2}\delta(t).$$

In general, we can define the product of two signals without any difficulty, as long as the signals do not contain singularities (discontinuities, impulses, or the other singularities introduced in Section 2.5) whose locations coincide. When the locations do coincide, the product is undefined. As an example, show that the signal

$$g(t) = \int_{-\infty}^{+\infty} u(\tau)\delta(t-\tau)d\tau$$

is identical to $u(t)$; that is, it is 0 for $t < 0$, it equals 1 for $t > 0$, and it is undefined for $t = 0$.

1.40. **(a)** Show that if a system is *either* additive or homogeneous, it has the property that if the input is identically zero, then the output is also identically zero.

(b) Determine a system (either in continuous or discrete time) that is *neither* additive *nor* homogeneous but which has a zero output if the input is identically zero.

(c) From part (a), can you conclude that if the input to a linear system is zero between times t_1 and t_2 in continuous time or between times n_1 and n_2 in discrete time, then its output must also be zero between these same times? Explain your answer.

1.41. Consider a system S with input $x[n]$ and output $y[n]$ related by

$$y[n] = x[n]\{g[n] + g[n-1]\}.$$

(a) If $g[n] = 1$ for all n, show that S is time invariant.
(b) If $g[n] = n$, show that S is not time invariant.
(c) If $g[n] = 1 + (-1)^n$, show that S is time invariant.

1.42. **(a)** Is the following statement true or false?

The series interconnection of two linear time-invariant systems is itself a linear, time-invariant system.

Justify your answer.

(b) Is the following statement true or false?

The series interconnection of two nonlinear systems is itself nonlinear.

Justify your answer.

(c) Consider three systems with the following input-output relationships:

$$\text{System 1:} \qquad y[n] = \begin{cases} x[n/2], & n \text{ even} \\ 0, & n \text{ odd} \end{cases},$$

$$\text{System 2:} \quad y[n] = x[n] + \frac{1}{2}x[n-1] + \frac{1}{4}x[n-2],$$

$$\text{System 3:} \quad y[n] = x[2n].$$

Suppose that these systems are connected in series as depicted in Figure P1.42. Find the input-output relationship for the overall interconnected system. Is this system linear? Is it time invariant?

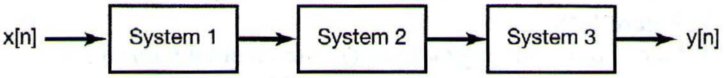


Figure P1.42

1.43. **(a)** Consider a time-invariant system with input $x(t)$ and output $y(t)$. Show that if $x(t)$ is periodic with period T, then so is $y(t)$. Show that the analogous result also holds in discrete time.

(b) Give an example of a time-invariant system and a nonperiodic input signal $x(t)$ such that the corresponding output $y(t)$ is periodic.

1.44. **(a)** Show that causality for a continuous-time linear system is equivalent to the following statement:

For any time t_0 and any input $x(t)$ such that $x(t) = 0$ for $t < t_0$, the corresponding output $y(t)$ must also be zero for $t < t_0$.

The analogous statement can be made for a discrete-time linear system.

(b) Find a nonlinear system that satisfies the foregoing condition but is not causal.

(c) Find a nonlinear system that is causal but does not satisfy the condition.

(d) Show that invertibility for a discrete-time linear system is equivalent to the following statement:

The only input that produces $y[n] = 0$ for all n is $x[n] = 0$ for all n.

The analogous statement is also true for a continuous-time linear system.

(e) Find a nonlinear system that satisfies the condition of part (d) but is not invertible.

1.45. In Problem 1.37, we introduced the concept of correlation functions. It is often important in practice to compute the correlation function $\phi_{hx}(t)$, where $h(t)$ is a fixed given signal, but where $x(t)$ may be any of a wide variety of signals. In this case, what is done is to design a system S with input $x(t)$ and output $\phi_{hx}(t)$.

(a) Is S linear? Is S time invariant? Is S causal? Explain your answers.

(b) Do any of your answers to part (a) change if we take as the output $\phi_{xh}(t)$ rather than $\phi_{hx}(t)$?

1.46. Consider the feedback system of Figure P1.46. Assume that $y[n] = 0$ for $n < 0$.

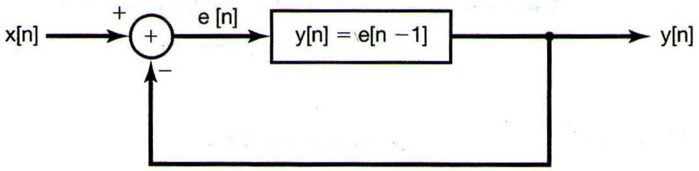


Figure P1.46

(a) Sketch the output when $x[n] = \delta[n]$.
(b) Sketch the output when $x[n] = u[n]$.

1.47. **(a)** Let S denote an incrementally linear system, and let $x_1[n]$ be an arbitrary input signal to S with corresponding output $y_1[n]$. Consider the system illustrated in Figure P1.47(a). Show that this system is linear and that, in fact, the overall input-output relationship between $x[n]$ and $y[n]$ does not depend on the particular choice of $x_1[n]$.

(b) Use the result of part (a) to show that S can be represented in the form shown in Figure 1.48.

(c) Which of the following systems are incrementally linear? Justify your answers, and if a system is incrementally linear, identify the linear system L and the zero-input response $y_0[n]$ or $y_0(t)$ for the representation of the system as shown in Figure 1.48.

(i) $y[n] = n + x[n] + 2x[n+4]$

(ii) $$y[n] = \begin{cases} n/2, & n \text{ even} \\ (n-1)/2 + \displaystyle\sum_{k=-\infty}^{(n-1)/2} x[k], & n \text{ odd} \end{cases}$$

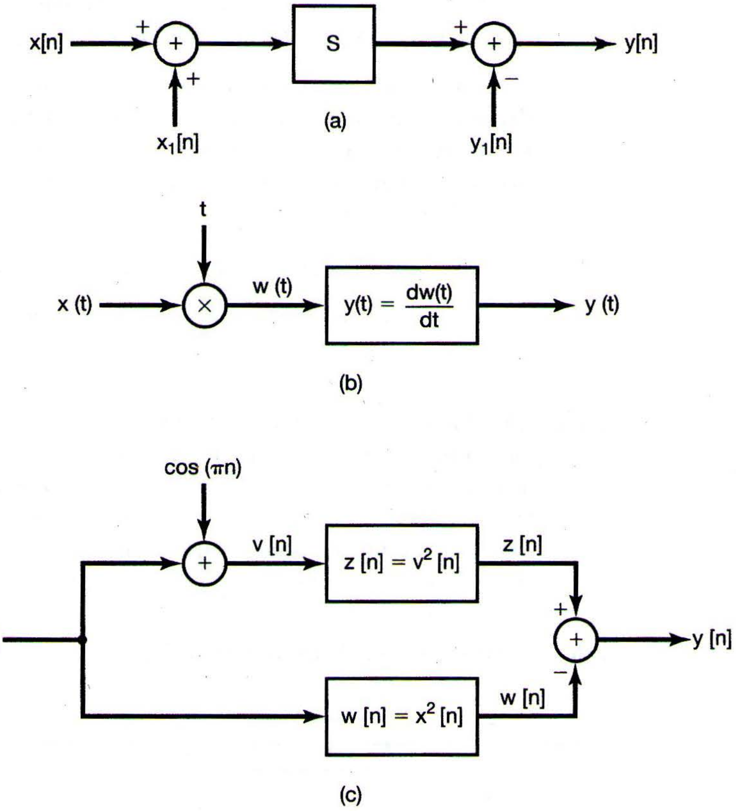


Figure P1.47

(iii) $y[n] = \begin{cases} x[n] - x[n-1] + 3, & \text{if } x[0] \geq 0 \\ x[n] - x[n-1] - 3, & \text{if } x[0] < 0 \end{cases}$

(iv) The system depicted in Figure P1.47(b).

(v) The system depicted in Figure P1.47(c).

(d) Suppose that a particular incrementally linear system has a representation as in Figure 1.48, with L denoting the linear system and $y_0[n]$ the zero-input response. Show that S is time invariant if and only if L is a time-invariant system *and* $y_0[n]$ is constant.

MATHEMATICAL REVIEW

The complex number z can be expressed in several ways. The *Cartesian* or *rectangular* form for z is

$$z = x + jy,$$

where $j = \sqrt{-1}$ and x and y are real numbers referred to respectively as the *real part* and the *imaginary part* of z. As we indicated earlier, we will often use the notation

$$x = \mathcal{R}e\{z\},\ y = \mathcal{I}m\{z\}.$$

The complex number z can also be represented in *polar form* as

$$z = re^{j\theta},$$

where $r > 0$ is the *magnitude* of z and θ is the *angle* or *phase* of z. These quantities will often be written as

$$r = |z|,\ \theta = \sphericalangle z.$$

The relationship between these two representations of complex numbers can be determined either from *Euler's relation*,

$$e^{j\theta} = \cos\theta + j\sin\theta,$$

or by plotting z in the complex plane, as shown in Figure P1.48, in which the coordinate axes are $\mathcal{R}e\{z\}$ along the horizontal axis and $\mathcal{I}m\{z\}$ along the vertical axis. With respect to this graphical representation, x and y are the Cartesian coordinates of z, and r and θ are its polar coordinates.

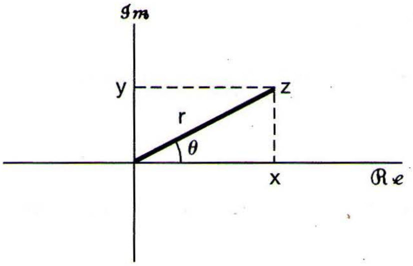


Figure P1.48

1.48. Let z_0 be a complex number with polar coordinates (r_0, θ_0) and Cartesian coordinates (x_0, y_0). Determine expressions for the Cartesian coordinates of the following complex numbers in terms of x_0 and y_0. Plot the points z_0, z_1, z_2, z_3, z_4, and z_5 in the complex plane when $r_0 = 2$ and $\theta_0 = \pi/4$ and when $r_0 = 2$ and $\theta_0 = \pi/2$. Indicate on your plots the real and imaginary parts of each point.

(a) $z_1 = r_0 e^{-j\theta_0}$ **(b)** $z_2 = r_0$ **(c)** $z_3 = r_0 e^{j(\theta_0+\pi)}$
(d) $z_4 = r_0 e^{j(-\theta_0+\pi)}$ **(e)** $z_5 = r_0 e^{j(\theta_0+2\pi)}$

1.49. Express each of the following complex numbers in polar form, and plot them in the complex plane, indicating the magnitude and angle of each number:

(a) $1 + j\sqrt{3}$ **(b)** -5 **(c)** $-5 - 5j$
(d) $3 + 4j$ **(e)** $(1 - j\sqrt{3})^3$ **(f)** $(1 + j)^5$
(g) $(\sqrt{3} + j^3)(1 - j)$ **(h)** $\frac{2-j(6/\sqrt{3})}{2+j(6/\sqrt{3})}$ **(i)** $\frac{1+j\sqrt{3}}{\sqrt{3}+j}$
(j) $j(1 + j)e^{j\pi/6}$ **(k)** $(\sqrt{3} + j)2\sqrt{2}e^{-j\pi/4}$ **(l)** $\frac{e^{j\pi/3}-1}{1+j\sqrt{3}}$

1.50. **(a)** Using Euler's relationship or Figure P1.48, determine expressions for x and y in terms of r and θ.

(b) Determine expressions for r and θ in terms of x and y.

(c) If we are given only r and $\tan\theta$, can we uniquely determine x and y? Explain your answer.

1.51. Using Euler's relation, derive the following relationships:

(a) $\cos\theta = \frac{1}{2}(e^{j\theta} + e^{-j\theta})$
(b) $\sin\theta = \frac{1}{2j}(e^{j\theta} - e^{-j\theta})$
(c) $\cos^2\theta = \frac{1}{2}(1 + \cos 2\theta)$
(d) $(\sin\theta)(\sin\phi) = \frac{1}{2}\cos(\theta - \phi) - \frac{1}{2}\cos(\theta + \phi)$
(e) $\sin(\theta + \phi) = \sin\theta\cos\phi + \cos\theta\sin\phi$

1.52. Let z denote a complex variable; that is,

$$z = x + jy = re^{j\theta}.$$

The *complex conjugate* of z is

$$z^* = x - jy = re^{-j\theta}.$$

Derive each of the following relations, where z, z_1, and z_2 are arbitrary complex numbers:

(a) $zz^* = r^2$
(b) $\frac{z}{z^*} = e^{j2\theta}$
(c) $z + z^* = 2\Re e\{z\}$
(d) $z - z^* = 2j\Im m\{z\}$
(e) $(z_1 + z_2)^* = z_1^* + z_2^*$
(f) $(az_1z_2)^* = az_1^*z_2^*$, where a is any real number
(g) $(\frac{z_1}{z_2})^* = \frac{z_1^*}{z_2^*}$
(h) $\Re e\{\frac{z_1}{z_2}\} = \frac{1}{2}[\frac{z_1z_2^*+z_1^*z_2}{z_2z_2^*}]$

1.53. Derive the following relations, where z, z_1, and z_2 are arbitrary complex numbers:

(a) $(e^z)^* = e^{z^*}$
(b) $z_1z_2^* + z_1^*z_2 = 2\Re e\{z_1z_2^*\} = 2\Re e\{z_1^*z_2\}$

(c) $|z| = |z^*|$
(d) $|z_1 z_2| = |z_1||z_2|$
(e) $\mathcal{R}e\{z\} \leq |z|, \mathcal{I}m\{z\} \leq |z|$
(f) $|z_1 z_2^* + z_1^* z_2| \leq 2|z_1 z_2|$
(g) $(|z_1| - |z_2|)^2 \leq |z_1 + z_2|^2 \leq (|z_1| + |z_2|)^2$

1.54. The relations considered in this problem are used on many occasions throughout the book.

(a) Prove the validity of the following expression:

$$\sum_{n=0}^{N-1} \alpha^n = \begin{cases} N, & \alpha = 1 \\ \frac{1-\alpha^N}{1-\alpha}, & \text{for any complex number } \alpha \neq 1 \end{cases}.$$

This is often referred to as the *finite sum formula.*

(b) Show that if $|\alpha| < 1$, then

$$\sum_{n=0}^{\infty} \alpha^n = \frac{1}{1-\alpha}.$$

This is often referred to as the *infinite sum formula.*

(c) Show also if $|\alpha| < 1$, then

$$\sum_{n=0}^{\infty} n\alpha^n = \frac{\alpha}{(1-\alpha)^2}.$$

(d) Evaluate

$$\sum_{n=k}^{\infty} \alpha^n,$$

assuming that $|\alpha| < 1$.

1.55. Using the results from Problem 1.54, evaluate each of the following sums and express your answer in Cartesian (rectangular) form:

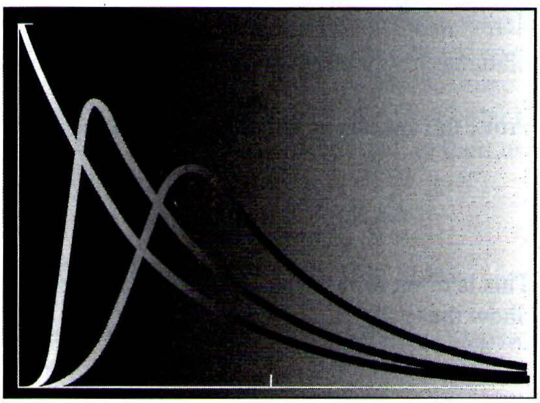
(a) $\sum_{n=0}^{9} e^{j\pi n/2}$ **(b)** $\sum_{n=-2}^{7} e^{j\pi n/2}$
(c) $\sum_{n=0}^{\infty} (\frac{1}{2})^n e^{j\pi n/2}$ **(d)** $\sum_{n=2}^{\infty} (\frac{1}{2})^n e^{j\pi n/2}$
(e) $\sum_{n=0}^{9} \cos(\frac{\pi}{2} n)$ **(f)** $\sum_{n=0}^{\infty} (\frac{1}{2})^n \cos(\frac{\pi}{2} n)$

1.56. Evaluate each of the following integrals, and express your answer in Cartesian (rectangular) form:

(a) $\int_0^4 e^{j\pi t/2} dt$ **(b)** $\int_0^6 e^{j\pi t/2} dt$
(c) $\int_2^8 e^{j\pi t/2} dt$ **(d)** $\int_0^{\infty} e^{-(1+j)t} dt$
(e) $\int_0^{\infty} e^{-t} \cos(t) dt$ **(f)** $\int_0^{\infty} e^{-2t} \sin(3t) dt$

2

LINEAR TIME-INVARIANT SYSTEMS



2.0 INTRODUCTION

In Section 1.6 we introduced and discussed a number of basic system properties. Two of these, linearity and time invariance, play a fundamental role in signal and system analysis for two major reasons. First, many physical processes possess these properties and thus can be modeled as linear time-invariant (LTI) systems. In addition, LTI systems can be analyzed in considerable detail, providing both insight into their properties and a set of powerful tools that form the core of signal and system analysis.

A principal objective of this book is to develop an understanding of these properties and tools and to provide an introduction to several of the very important applications in which the tools are used. In this chapter, we begin the development by deriving and examining a fundamental and extremely useful representation for LTI systems and by introducing an important class of these systems.

One of the primary reasons LTI systems are amenable to analysis is that any such system possesses the superposition property described in Section 1.6.6. As a consequence, if we can represent the input to an LTI system in terms of a linear combination of a set of basic signals, we can then use superposition to compute the output of the system in terms of its responses to these basic signals.

As we will see in the following sections, one of the important characteristics of the unit impulse, both in discrete time and in continuous time, is that very general signals can be represented as linear combinations of delayed impulses. This fact, together with the properties of superposition and time invariance, will allow us to develop a complete characterization of any LTI system in terms of its response to a unit impulse. Such a representation, referred to as the convolution sum in the discrete-time case and the convolution integral in continuous time, provides considerable analytical convenience in dealing with LTI systems. Following our development of the convolution sum and the convolution integral we use these characterizations to examine some of the other properties of LTI systems. We then consider the class of continuous-time systems described by linear constant-coefficient differential equations and its discrete-time counterpart, the class of systems described by linear constant-coefficient difference equations. We will return to examine these two very important classes of systems on a number of occasions in subsequent chapters. Finally, we will take another look at the continuous-time unit impulse function and a number of other signals that are closely related to it in order to provide some additional insight into these idealized signals and, in particular, to their use and interpretation in the context of analyzing LTI systems.

2.1 DISCRETE-TIME LTI SYSTEMS: THE CONVOLUTION SUM

2.1.1 The Representation of Discrete-Time Signals in Terms of Impulses

The key idea in visualizing how the discrete-time unit impulse can be used to construct any discrete-time signal is to think of a discrete-time signal as a sequence of individual impulses. To see how this intuitive picture can be turned into a mathematical representation, consider the signal $x[n]$ depicted in Figure 2.1(a). In the remaining parts of this figure, we have depicted five time-shifted, scaled unit impulse sequences, where the scaling on each impulse equals the value of $x[n]$ at the particular instant the unit sample occurs. For example,

$$x[-1]\delta[n+1] = \begin{cases} x[-1], & n = -1 \\ 0, & n \neq -1 \end{cases},$$

$$x[0]\delta[n] = \begin{cases} x[0], & n = 0 \\ 0, & n \neq 0 \end{cases},$$

$$x[1]\delta[n-1] = \begin{cases} x[1], & n = 1 \\ 0, & n \neq 1 \end{cases}.$$

Therefore, the sum of the five sequences in the figure equals $x[n]$ for $-2 \leq n \leq 2$. More generally, by including additional shifted, scaled impulses, we can write

$$\begin{aligned} x[n] = \ldots &+ x[-3]\delta[n+3] + x[-2]\delta[n+2] + x[-1]\delta[n+1] + x[0]\delta[n] \\ &+ x[1]\delta[n-1] + x[2]\delta[n-2] + x[3]\delta[n-3] + \ldots . \end{aligned} \tag{2.1}$$

For any value of n, only one of the terms on the right-hand side of eq. (2.1) is nonzero, and the scaling associated with that term is precisely $x[n]$. Writing this summation in a more compact form, we have

$$x[n] = \sum_{k=-\infty}^{+\infty} x[k]\delta[n-k]. \tag{2.2}$$

This corresponds to the representation of an arbitrary sequence as a linear combination of shifted unit impulses $\delta[n-k]$, where the weights in this linear combination are $x[k]$. As an example, consider $x[n] = u[n]$, the unit step. In this case, since $u[k] = 0$ for $k < 0$

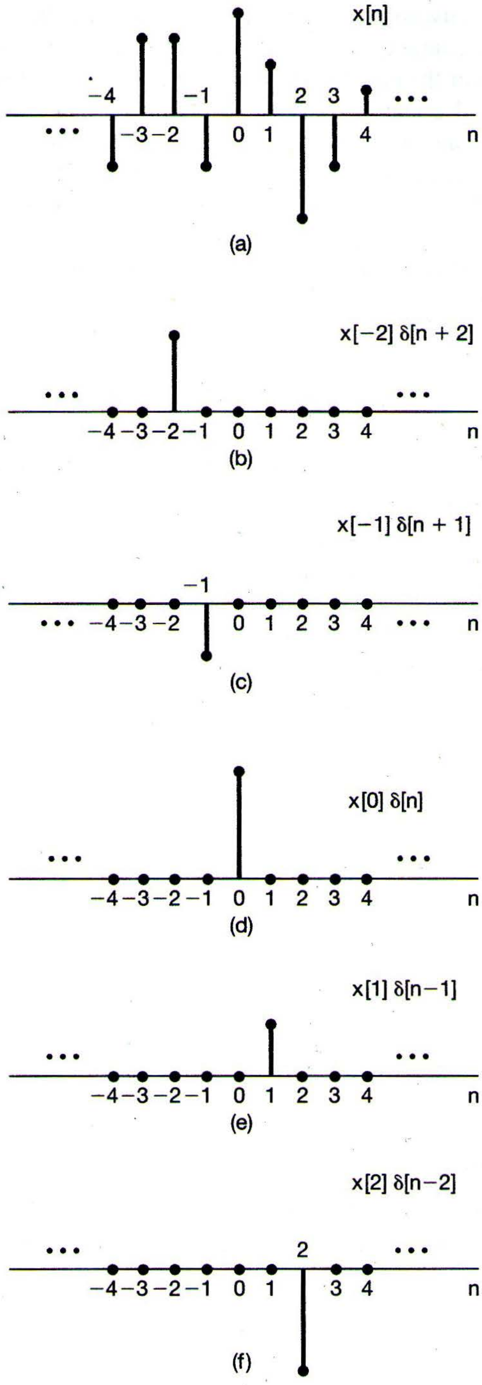

x[n]
−4
−3
−2
−1
0
1
2
3
4
n
(a)
x[−2] δ[n + 2]
−4 −3 −2 −1 0 1 2 3 4
n
(b)
x[−1] δ[n + 1]
−1
−4 −3 −2 0 1 2 3 4
n
(c)
x[0] δ[n]
−4 −3 −2 −1 0 1 2 3 4
n
(d)
x[1] δ[n−1]
−4 −3 −2 −1 0 1 2 3 4
n
(e)
x[2] δ[n−2]
2
−4 −3 −2 −1 0 1 3 4
n
(f)

Figure 2.1 Decomposition of a discrete-time signal into a weighted sum of shifted impulses.

and $u[k] = 1$ for $k \geq 0$, eq. (2.2) becomes

$$u[n] = \sum_{k=0}^{+\infty} \delta[n-k],$$

which is identical to the expression we derived in Section 1.4. [See eq. (1.67).]

Equation (2.2) is called the *sifting property* of the discrete-time unit impulse. Because the sequence $\delta[n-k]$ is nonzero only when $k = n$, the summation on the right-hand side of eq. (2.2) "sifts" through the sequence of values $x[k]$ and preserves only the value corresponding to $k = n$. In the next subsection, we will exploit this representation of discrete-time signals in order to develop the convolution-sum representation for a discrete-time LTI system.

2.1.2 The Discrete-Time Unit Impulse Response and the Convolution-Sum Representation of LTI Systems

The importance of the sifting property of eqs. (2.1) and (2.2) lies in the fact that it represents $x[n]$ as a superposition of scaled versions of a very simple set of elementary functions, namely, shifted unit impulses $\delta[n-k]$, each of which is nonzero (with value 1) at a single point in time specified by the corresponding value of k. The response of a linear system to $x[n]$ will be the superposition of the scaled responses of the system to each of these shifted impulses. Moreover, the property of time invariance tells us that the responses of a time-invariant system to the time-shifted unit impulses are simply time-shifted versions of one another. The convolution-sum representation for discrete-time systems that are *both* linear and time invariant results from putting these two basic facts together.

More specifically, consider the response of a linear (but possibly time-varying) system to an arbitrary input $x[n]$. We can represent the input through eq. (2.2) as a linear combination of shifted unit impulses. Let $h_k[n]$ denote the response of the linear system to the shifted unit impulse $\delta[n-k]$. Then, from the superposition property for a linear system [eqs. (1.123) and (1.124)], the response $y[n]$ of the linear system to the input $x[n]$ in eq. (2.2) is simply the weighted linear combination of these basic responses. That is, with the input $x[n]$ to a linear system expressed in the form of eq. (2.2), the output $y[n]$ can be expressed as

$$y[n] = \sum_{k=-\infty}^{+\infty} x[k]h_k[n]. \tag{2.3}$$

Thus, according to eq. (2.3), if we know the response of a linear system to the set of shifted unit impulses, we can construct the response to an arbitrary input. An interpretation of eq. (2.3) is illustrated in Figure 2.2. The signal $x[n]$ is applied as the input to a linear system whose responses $h_{-1}[n]$, $h_0[n]$, and $h_1[n]$ to the signals $\delta[n+1]$, $\delta[n]$, and $\delta[n-1]$, respectively, are depicted in Figure 2.2(b). Since $x[n]$ can be written as a linear combination of $\delta[n+1]$, $\delta[n]$, and $\delta[n-1]$, superposition allows us to write the response to $x[n]$ as a linear combination of the responses to the individual shifted impulses. The individual shifted and scaled impulses that constitute $x[n]$ are illustrated on the left-hand side of Figure 2.2(c), while the responses to these component signals are pictured on the right-hand side. In Figure 2.2(d) we have depicted the actual input $x[n]$, which is the sum of the components on the left side of Figure 2.2(c) and the actual output $y[n]$, which, by

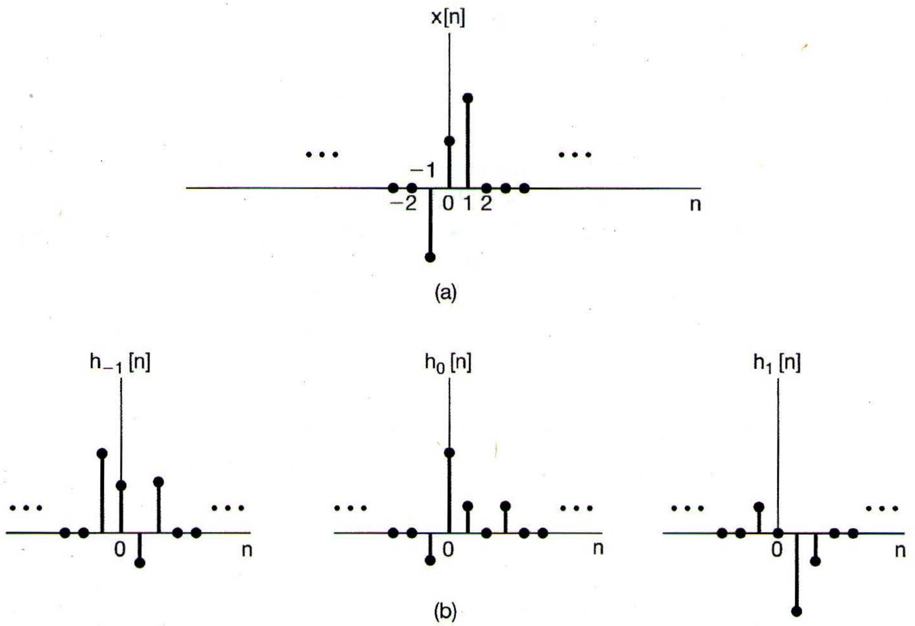


Figure 2.2 Graphical interpretation of the response of a discrete-time linear system as expressed in eq. (2.3).

superposition, is the sum of the components on the right side of Figure 2.2(c). Thus, the response at time n of a linear system is simply the superposition of the responses due to the input value at each point in time.

In general, of course, the responses $h_k[n]$ need not be related to each other for different values of k. However, if the linear system is also *time invariant*, then these responses to time-shifted unit impulses are all time-shifted versions of each other. Specifically, since $\delta[n-k]$ is a time-shifted version of $\delta[n]$, the response $h_k[n]$ is a time-shifted version of $h_0[n]$; i.e.,

$$h_k[n] = h_0[n-k]. \tag{2.4}$$

For notational convenience, we will drop the subscript on $h_0[n]$ and define the *unit impulse (sample) response*

$$h[n] = h_0[n]. \tag{2.5}$$

That is, $h[n]$ is the output of the LTI system when $\delta[n]$ is the input. Then for an LTI system, eq. (2.3) becomes

$$\boxed{y[n] = \sum_{k=-\infty}^{+\infty} x[k]h[n-k].} \tag{2.6}$$

This result is referred to as the *convolution sum* or *superposition sum*, and the operation on the right-hand side of eq. (2.6) is known as the *convolution* of the sequences $x[n]$ and $h[n]$. We will represent the operation of convolution symbolically as

$$y[n] = x[n] * h[n]. \tag{2.7}$$

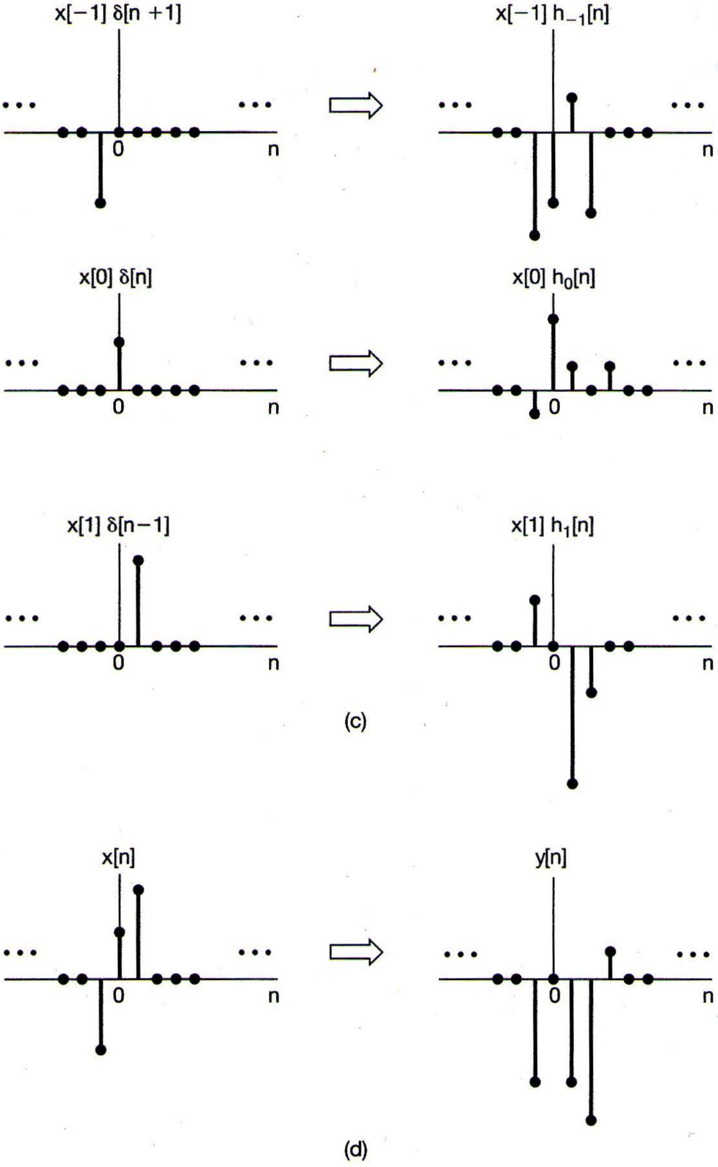


Figure 2.2 *Continued*

Note that eq. (2.6) expresses the response of an LTI system to an arbitrary input in terms of the system's response to the unit impulse. From this, we see that an LTI system is completely characterized by its response to a single signal, namely, its response to the unit impulse.

The interpretation of eq. (2.6) is similar to the one we gave for eq. (2.3), where, in the case of an LTI system, the response due to the input $x[k]$ applied at time k is $x[k]h[n-k]$; i.e., it is a shifted and scaled version (an "echo") of $h[n]$. As before, the actual output is the superposition of all these responses.

Example 2.1

Consider an LTI system with impulse response $h[n]$ and input $x[n]$, as illustrated in Figure 2.3(a). For this case, since only $x[0]$ and $x[1]$ are nonzero, eq. (2.6) simplifies to the expression

$$y[n] = x[0]h[n-0] + x[1]h[n-1] = 0.5h[n] + 2h[n-1]. \tag{2.8}$$

The sequences $0.5h[n]$ and $2h[n-1]$ are the two echoes of the impulse response needed for the superposition involved in generating $y[n]$. These echoes are displayed in Figure 2.3(b). By summing the two echoes for each value of n, we obtain $y[n]$, which is shown in Figure 2.3(c).

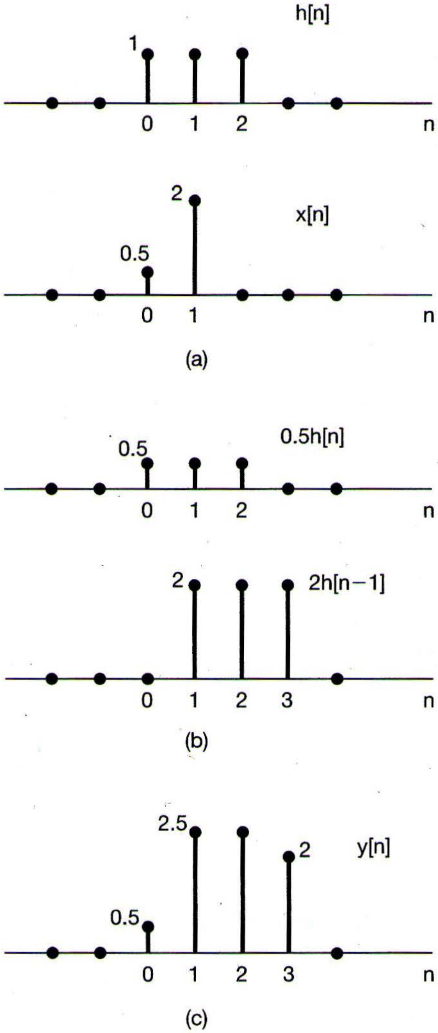


Figure 2.3 (a) The impulse response $h[n]$ of an LTI system and an input $x[n]$ to the system; (b) the responses or "echoes," $0.5h[n]$ and $2h[n-1]$, to the nonzero values of the input, namely, $x[0] = 0.5$ and $x[1] = 2$; (c) the overall response $y[n]$, which is the sum of the echos in (b).

By considering the effect of the superposition sum on each individual output sample, we obtain another very useful way to visualize the calculation of $y[n]$ using the convolution sum. In particular, consider the evaluation of the output value at some specific time n. A particularly convenient way of displaying this calculation graphically begins with the two signals $x[k]$ and $h[n-k]$ viewed as functions of k. Multiplying these two functions, we obtain a sequence $g[k] = x[k]h[n-k]$, which, at each time k, is seen to represent the contribution of $x[k]$ to the output at time n. We conclude that summing all the samples in the sequence of $g[k]$ yields the output value at the selected time n. Thus, to calculate $y[n]$ for all values of n requires repeating this procedure for each value of n. Fortunately, changing the value of n has a very simple graphical interpretation for the two signals $x[k]$ and $h[n-k]$, viewed as functions of k. The following examples illustrate this and the use of the aforementioned viewpoint in evaluating convolution sums.

Example 2.2

Let us consider again the convolution problem encountered in Example 2.1. The sequence $x[k]$ is shown in Figure 2.4(a), while the sequence $h[n-k]$, for n fixed and viewed as a function of k, is shown in Figure 2.4(b) for several different values of n. In sketching these sequences, we have used the fact that $h[n-k]$ (viewed as a function of k with n fixed) is a time-reversed and shifted version of the impulse response $h[k]$. In particular, as k increases, the argument $n-k$ decreases, explaining the need to perform a time reversal of $h[k]$. Knowing this, then in order to sketch the signal $h[n-k]$, we need only determine its value for some particular value of k. For example, the argument $n-k$ will equal 0 at the value $k = n$. Thus, if we sketch the signal $h[-k]$, we can obtain the signal $h[n-k]$ simply by shifting to the right (by n) if n is positive or to the left if n is negative. The result for our example for values of $n < 0$, $n = 0, 1, 2, 3$, and $n > 3$ are shown in Figure 2.4(b).

Having sketched $x[k]$ and $h[n-k]$ for any particular value of n, we multiply these two signals and sum over all values of k. For our example, for $n < 0$, we see from Figure 2.4 that $x[k]h[n-k] = 0$ for all k, since the nonzero values of $x[k]$ and $h[n-k]$ do not overlap. Consequently, $y[n] = 0$ for $n < 0$. For $n = 0$, since the product of the sequence $x[k]$ with the sequence $h[0-k]$ has only one nonzero sample with the value 0.5, we conclude that

$$y[0] = \sum_{k=-\infty}^{\infty} x[k]h[0-k] = 0.5. \tag{2.9}$$

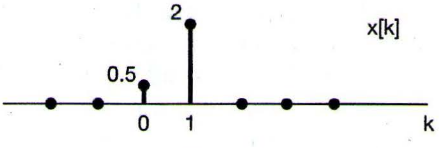
The product of the sequence $x[k]$ with the sequence $h[1-k]$ has two nonzero samples, which may be summed to obtain

$$y[1] = \sum_{k=-\infty}^{\infty} x[k]h[1-k] = 0.5 + 2.0 = 2.5. \tag{2.10}$$

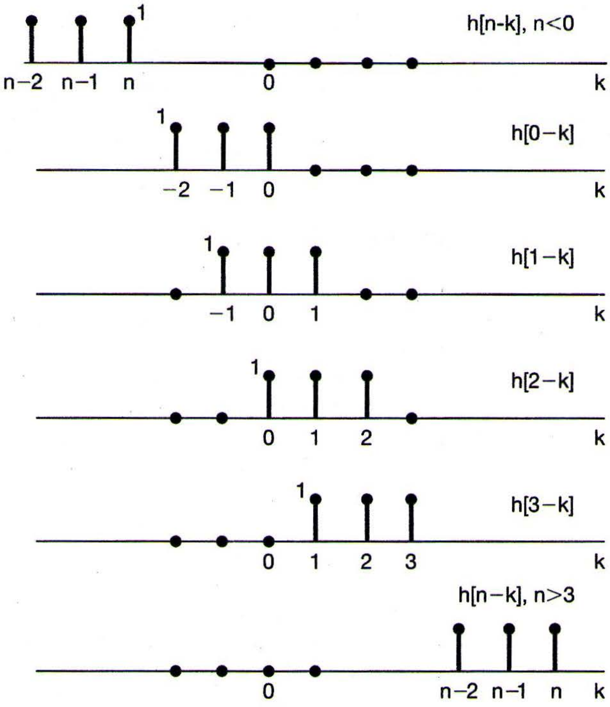
Similarly,

$$y[2] = \sum_{k=-\infty}^{\infty} x[k]h[2-k] = 0.5 + 2.0 = 2.5, \tag{2.11}$$

and



(a)



(b)

Figure 2.4 Interpretation of eq. (2.6) for the signals $h[n]$ and $x[n]$ in Figure 2.3; (a) the signal $x[k]$ and (b) the signal $h[n-k]$ (as a function of k with n fixed) for several values of n ($n < 0$; $n = 0, 1, 2, 3$; $n > 3$). Each of these signals is obtained by reflection and shifting of the unit impulse response $h[k]$. The response $y[n]$ for each value of n is obtained by multiplying the signals $x[k]$ and $h[n-k]$ in (a) and (b) and then summing the products over all values of k. The calculation for this example is carried out in detail in Example 2.2.

$$y[3] = \sum_{k=-\infty}^{\infty} x[k]h[3-k] = 2.0. \tag{2.12}$$

Finally, for $n > 3$, the product $x[k]h[n-k]$ is zero for all k, from which we conclude that $y[n] = 0$ for $n > 3$. The resulting output values agree with those obtained in Example 2.1.

Example 2.3

Consider an input $x[n]$ and a unit impulse response $h[n]$ given by

$$x[n] = \alpha^n u[n],$$
$$h[n] = u[n],$$

with $0 < \alpha < 1$. These signals are illustrated in Figure 2.5. Also, to help us in visualizing and calculating the convolution of the signals, in Figure 2.6 we have depicted the signal $x[k]$ followed by $h[-k]$, $h[-1-k]$, and $h[1-k]$ (that is, $h[n-k]$ for $n = 0, -1$, and $+1$) and, finally, $h[n-k]$ for an arbitrary positive value of n and an arbitrary negative value of n. From this figure, we note that for $n < 0$, there is no overlap between the nonzero points in $x[k]$ and $h[n-k]$. Thus, for $n < 0$, $x[k]h[n-k] = 0$ for all values of k, and hence, from eq. (2.6), we see that $y[n] = 0$, $n < 0$. For $n \geq 0$,

$$x[k]h[n-k] = \begin{cases} \alpha^k, & 0 \leq k \leq n \\ 0, & \text{otherwise} \end{cases}.$$

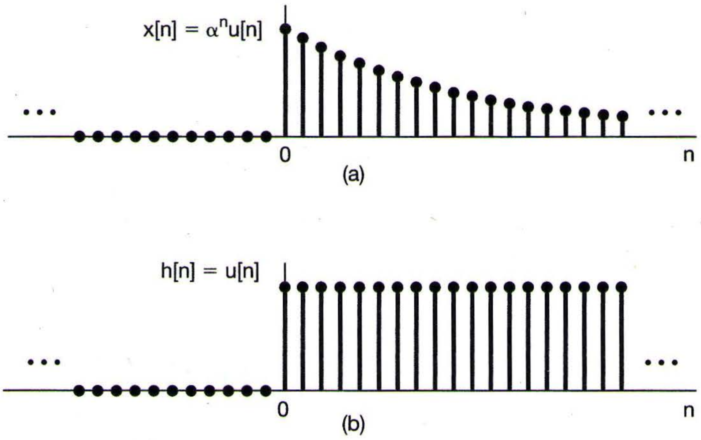


Figure 2.5 The signals $x[n]$ and $h[n]$ in Example 2.3.

Thus, for $n \geq 0$,

$$y[n] = \sum_{k=0}^{n} \alpha^k,$$

and using the result of Problem 1.54 we can write this as

$$y[n] = \sum_{k=0}^{n} \alpha^k = \frac{1-\alpha^{n+1}}{1-\alpha} \qquad \text{for } n \geq 0. \tag{2.13}$$

Thus, for all n,

$$y[n] = \left(\frac{1-\alpha^{n+1}}{1-\alpha}\right) u[n].$$

The signal $y[n]$ is sketched in Figure 2.7.

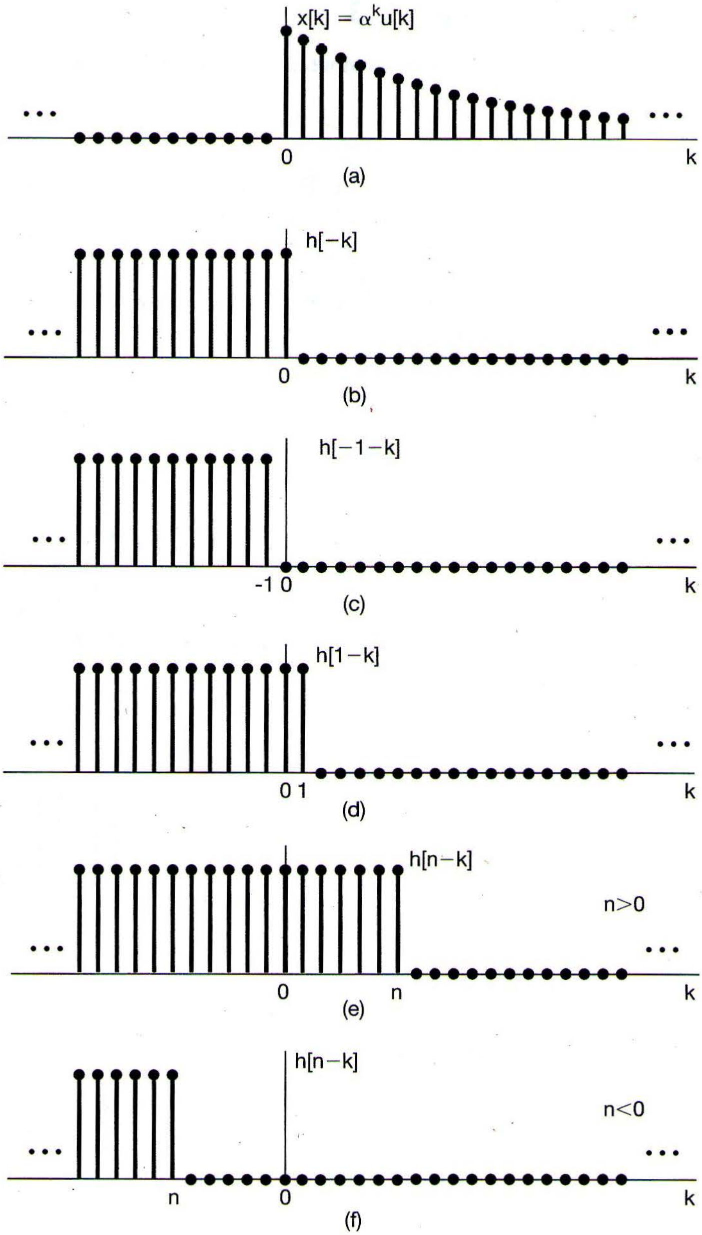

x[k] = αᵏu[k]
0
k
(a)
h[−k]
0
k
(b)
h[−1−k]
-1 0
k
(c)
h[1−k]
0 1
k
(d)
h[n−k]
n>0
0
n
k
(e)
h[n−k]
n<0
n
0
k
(f)

Figure 2.6 Graphical interpretation of the calculation of the convolution sum for Example 2.3.

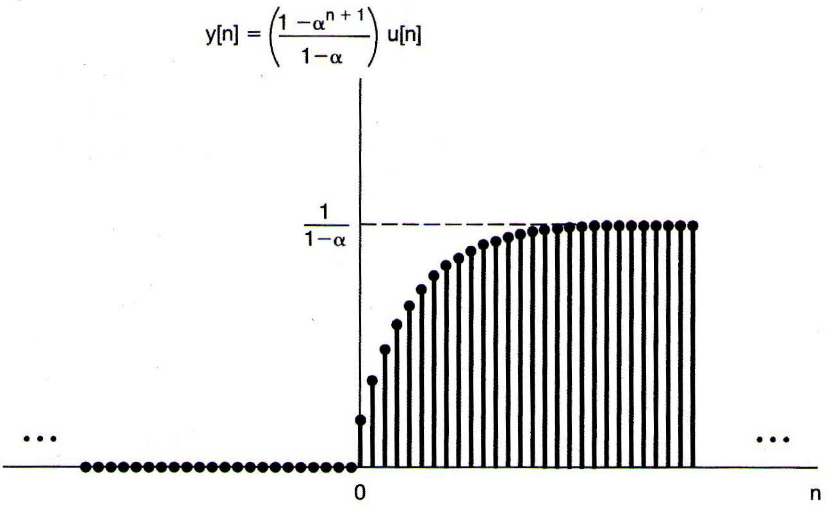


Figure 2.7 Output for Example 2.3.

The operation of convolution is sometimes described in terms of "sliding" the sequence $h[n-k]$ past $x[k]$. For example, suppose we have evaluated $y[n]$ for some particular value of n, say, $n = n_0$. That is, we have sketched the signal $h[n_0 - k]$, multiplied it by the signal $x[k]$, and summed the result over all values of k. To evaluate $y[n]$ at the next value of n—i.e., $n = n_0 + 1$—we need to sketch the signal $h[(n_0+1) - k]$. However, we can do this simply by taking the signal $h[n_0 - k]$ and shifting it to the right by one point. For each successive value of n, we continue this process of shifting $h[n-k]$ to the right by one point, multiplying by $x[k]$, and summing the result over k.

Example 2.4

As a further example, consider the two sequences

$$x[n] = \begin{cases} 1, & 0 \le n \le 4 \\ 0, & \text{otherwise} \end{cases}$$

and

$$h[n] = \begin{cases} \alpha^n, & 0 \le n \le 6 \\ 0, & \text{otherwise} \end{cases}.$$

These signals are depicted in Figure 2.8 for a positive value of $\alpha > 1$. In order to calculate the convolution of the two signals, it is convenient to consider five separate intervals for n. This is illustrated in Figure 2.9.

Interval 1. For $n < 0$, there is no overlap between the nonzero portions of $x[k]$ and $h[n-k]$, and consequently, $y[n] = 0$.

Interval 2. For $0 \le n \le 4$,

$$x[k]h[n-k] = \begin{cases} \alpha^{n-k}, & 0 \le k \le n \\ 0, & \text{otherwise} \end{cases}.$$

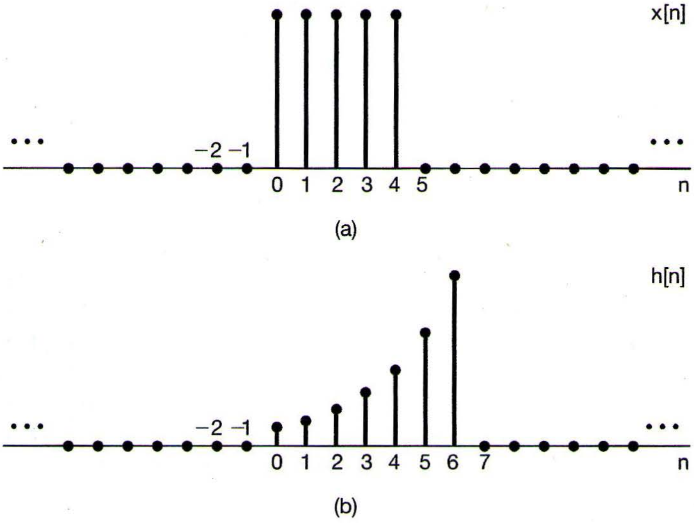


Figure 2.8 The signals to be convolved in Example 2.4.

Thus, in this interval,

$$y[n] = \sum_{k=0}^{n} \alpha^{n-k}. \tag{2.14}$$

We can evaluate this sum using the finite sum formula, eq. (2.13). Specifically, changing the variable of summation in eq. (2.14) from k to $r = n - k$, we obtain

$$y[n] = \sum_{r=0}^{n} \alpha^{r} = \frac{1 - \alpha^{n+1}}{1 - \alpha}.$$

Interval 3. For $n > 4$ but $n - 6 \leq 0$ (i.e., $4 < n \leq 6$),

$$x[k]h[n-k] = \begin{cases} \alpha^{n-k}, & 0 \leq k \leq 4 \\ 0, & \text{otherwise} \end{cases}.$$

Thus, in this interval,

$$y[n] = \sum_{k=0}^{4} \alpha^{n-k}. \tag{2.15}$$

Once again, we can use the geometric sum formula in eq. (2.13) to evaluate eq. (2.15). Specifically, factoring out the constant factor of α^n from the summation in eq. (2.15) yields

$$y[n] = \alpha^{n} \sum_{k=0}^{4} (\alpha^{-1})^{k} = \alpha^{n} \frac{1 - (\alpha^{-1})^{5}}{1 - \alpha^{-1}} = \frac{\alpha^{n-4} - \alpha^{n+1}}{1 - \alpha}. \tag{2.16}$$

Interval 4. For $n > 6$ but $n - 6 \leq 4$ (i.e., for $6 < n \leq 10$),

$$x[k]h[n-k] = \begin{cases} \alpha^{n-k}, & (n-6) \leq k \leq 4 \\ 0, & \text{otherwise} \end{cases},$$

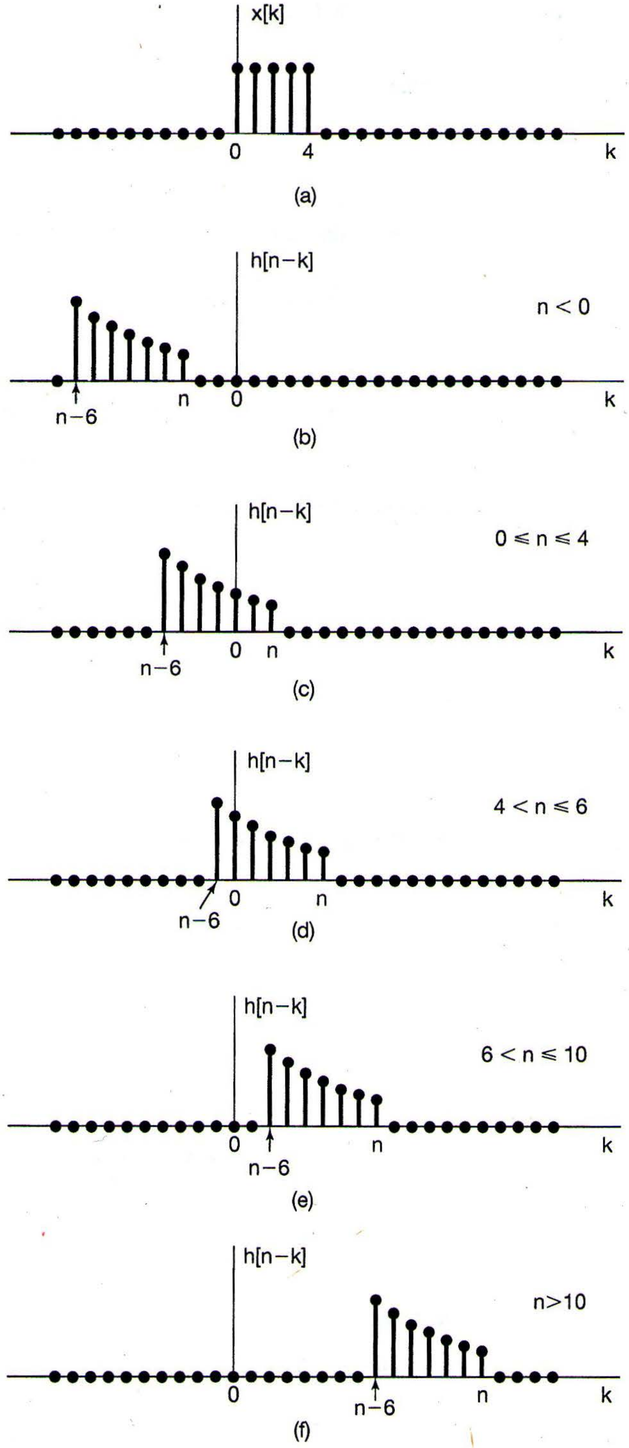

x[k]
0
4
k
(a)
h[n−k]
n < 0
n−6
n
0
k
(b)
h[n−k]
0 ≤ n ≤ 4
n−6
0
n
k
(c)
h[n−k]
4 < n ≤ 6
n−6
0
n
k
(d)
h[n−k]
6 < n ≤ 10
0
n−6
n
k
(e)
h[n−k]
n>10
0
n−6
n
k
(f)

Figure 2.9 Graphical interpretation of the convolution performed in Example 2.4.

so that

$$y[n] = \sum_{k=n-6}^{4} \alpha^{n-k}.$$

We can again use eq. (2.13) to evaluate this summation. Letting $r = k-n+6$, we obtain

$$y[n] = \sum_{r=0}^{10-n} \alpha^{6-r} = \alpha^6 \sum_{r=0}^{10-n} (\alpha^{-1})^r = \alpha^6 \frac{1-\alpha^{n-11}}{1-\alpha^{-1}} = \frac{\alpha^{n-4}-\alpha^7}{1-\alpha}.$$

Interval 5. For $n-6 > 4$, or equivalently, $n > 10$, there is no overlap between the nonzero portions of $x[k]$ and $h[n-k]$, and hence,

$$y[n] = 0.$$

Summarizing, then, we obtain

$$y[n] = \begin{cases} 0, & n < 0 \\ \dfrac{1-\alpha^{n+1}}{1-\alpha}, & 0 \le n \le 4 \\ \dfrac{\alpha^{n-4}-\alpha^{n+1}}{1-\alpha}, & 4 < n \le 6 \\ \dfrac{\alpha^{n-4}-\alpha^{7}}{1-\alpha}, & 6 < n \le 10 \\ 0, & 10 < n \end{cases},$$

which is pictured in Figure 2.10.

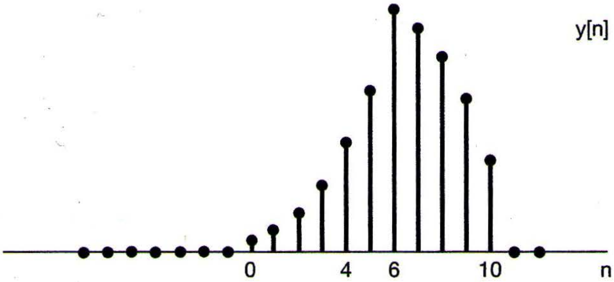


Figure 2.10 Result of performing the convolution in Example 2.4.

Example 2.5

Consider an LTI system with input $x[n]$ and unit impulse response $h[n]$ specified as follows:

$$x[n] = 2^n u[-n], \tag{2.17}$$

$$h[n] = u[n]. \tag{2.18}$$

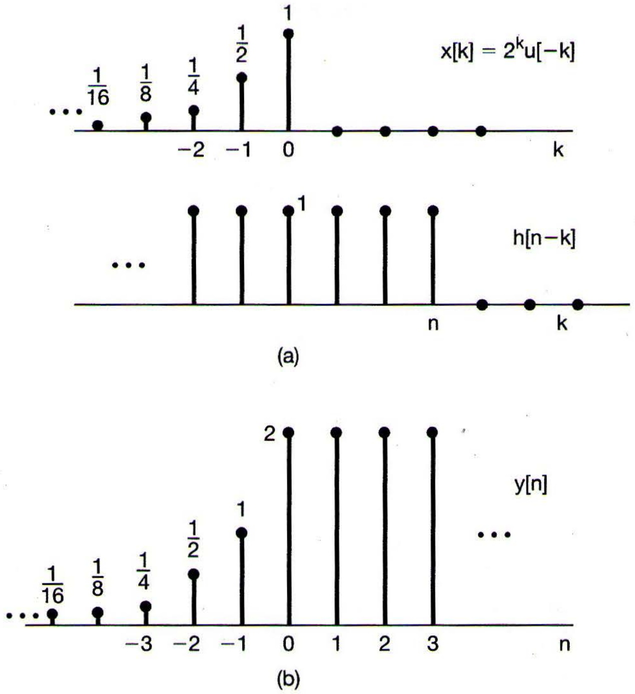


Figure 2.11 (a) The sequences $x[k]$ and $h[n-k]$ for the convolution problem considered in Example 2.5; (b) the resulting output signal $y[n]$.

The sequences $x[k]$ and $h[n-k]$ are plotted as functions of k in Figure 2.11(a). Note that $x[k]$ is zero for $k > 0$ and $h[n-k]$ is zero for $k > n$. We also observe that, regardless of the value of n, the sequence $x[k]h[n-k]$ always has nonzero samples along the k-axis. When $n \geq 0$, $x[k]h[n-k]$ has nonzero samples in the interval $k \leq 0$. It follows that, for $n \geq 0$,

$$y[n] = \sum_{k=-\infty}^{0} x[k]h[n-k] = \sum_{k=-\infty}^{0} 2^k. \tag{2.19}$$

To evaluate the infinite sum in eq. (2.19), we may use the *infinite sum formula*,

$$\sum_{k=0}^{\infty} \alpha^k = \frac{1}{1-\alpha}, \quad 0 < |\alpha| < 1. \tag{2.20}$$

Changing the variable of summation in eq. (2.19) from k to $r = -k$, we obtain

$$\sum_{k=-\infty}^{0} 2^k = \sum_{r=0}^{\infty} \left(\frac{1}{2}\right)^k = \frac{1}{1-(1/2)} = 2. \tag{2.21}$$

Thus, $y[n]$ takes on a constant value of 2 for $n \geq 0$.

When $n < 0$, $x[k]h[n-k]$ has nonzero samples for $k \leq n$. It follows that, for $n < 0$,

$$y[n] = \sum_{k=-\infty}^{n} x[k]h[n-k] = \sum_{k=-\infty}^{n} 2^k. \tag{2.22}$$

By performing a change of variable $l = -k$ and then $m = l + n$, we can again make use of the infinite sum formula, eq. (2.20), to evaluate the sum in eq. (2.22). The result is the following for $n < 0$:

$$y[n] = \sum_{l=-n}^{\infty} \left(\frac{1}{2}\right)^l = \sum_{m=0}^{\infty} \left(\frac{1}{2}\right)^{m-n} = \left(\frac{1}{2}\right)^{-n} \sum_{m=0}^{\infty} \left(\frac{1}{2}\right)^m = 2^n \cdot 2 = 2^{n+1}. \tag{2.23}$$

The complete sequence of $y[n]$ is sketched in Figure 2.11(b).

These examples illustrate the usefulness of visualizing the calculation of the convolution sum graphically. Moreover, in addition to providing a useful way in which to calculate the response of an LTI system, the convolution sum also provides an extremely useful representation for LTI systems that allows us to examine their properties in great detail. In particular, in Section 2.3 we will describe some of the properties of convolution and will also examine some of the system properties introduced in the previous chapter in order to see how these properties can be characterized for LTI systems.

2.2 CONTINUOUS-TIME LTI SYSTEMS: THE CONVOLUTION INTEGRAL

In analogy with the results derived and discussed in the preceding section, the goal of this section is to obtain a complete characterization of a continuous-time LTI system in terms of its unit impulse response. In discrete time, the key to our developing the convolution sum was the sifting property of the discrete-time unit impulse—that is, the mathematical representation of a signal as the superposition of scaled and shifted unit impulse functions. Intuitively, then, we can think of the discrete-time system as responding to a sequence of individual impulses. In continuous time, of course, we do not have a discrete sequence of input values. Nevertheless, as we discussed in Section 1.4.2, if we think of the unit impulse as the idealization of a pulse which is so short that its duration is inconsequential for any real, physical system, we can develop a representation for arbitrary continuous-time signals in terms of these idealized pulses with vanishingly small duration, or equivalently, impulses. This representation is developed in the next subsection, and, following that, we will proceed very much as in Section 2.1 to develop the convolution integral representation for continuous-time LTI systems.

2.2.1 The Representation of Continuous-Time Signals in Terms of Impulses

To develop the continuous-time counterpart of the discrete-time sifting property in eq. (2.2), we begin by considering a pulse or "staircase" approximation, $\hat{x}(t)$, to a continuous-time signal $x(t)$, as illustrated in Figure 2.12(a). In a manner similar to that

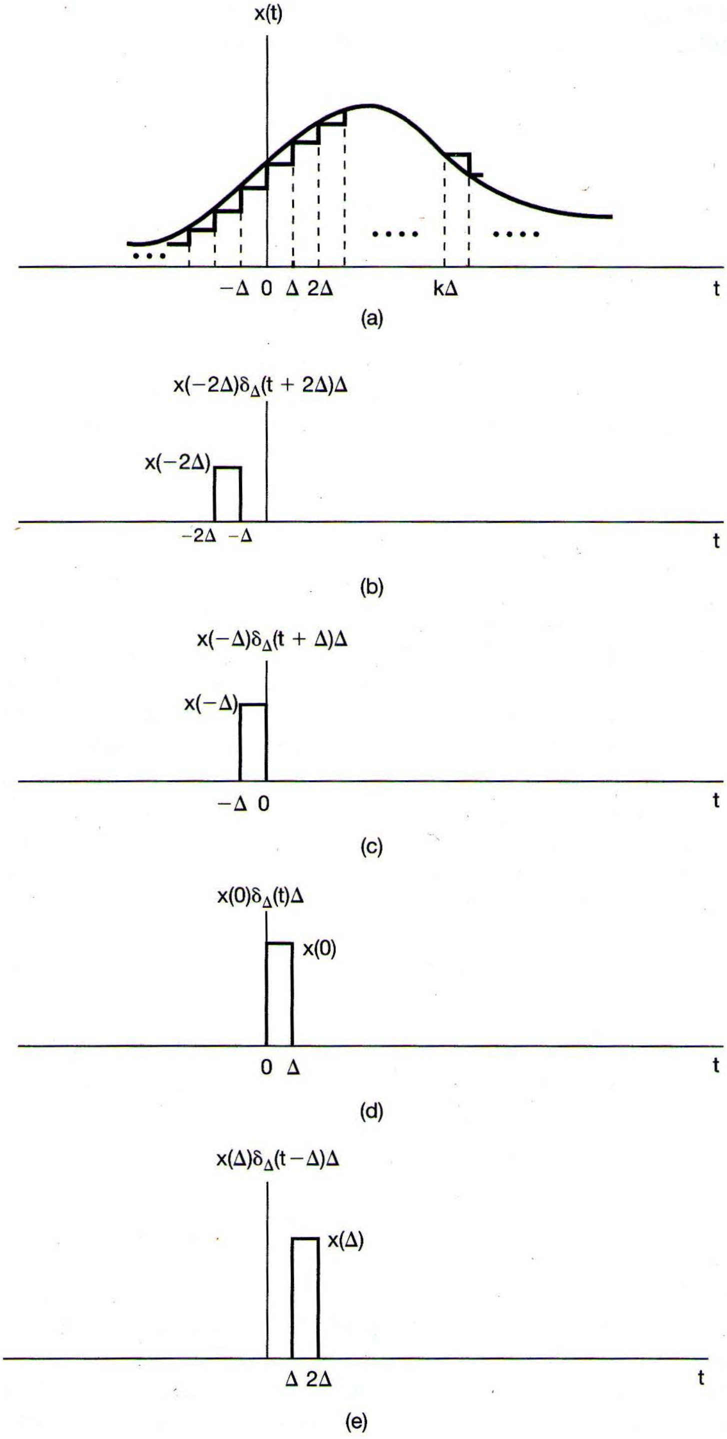

x(t)
−Δ 0 Δ 2Δ
kΔ
t
(a)
x(−2Δ)δ_Δ(t + 2Δ)Δ
x(−2Δ)
−2Δ −Δ
t
(b)
x(−Δ)δ_Δ(t + Δ)Δ
x(−Δ)
−Δ 0
t
(c)
x(0)δ_Δ(t)Δ
x(0)
0 Δ
t
(d)
x(Δ)δ_Δ(t − Δ)Δ
x(Δ)
Δ 2Δ
t
(e)

Figure 2.12 Staircase approximation to a continuous-time signal.

employed in the discrete-time case, this approximation can be expressed as a linear combination of delayed pulses, as illustrated in Figure 2.12(a)–(e). If we define

$$\delta_\Delta(t) = \begin{cases} \frac{1}{\Delta}, & 0 \le t < \Delta \\ 0, & \text{otherwise} \end{cases}, \tag{2.24}$$

then, since $\Delta\delta_\Delta(t)$ has unit amplitude, we have the expression

$$\hat{x}(t) = \sum_{k=-\infty}^{\infty} x(k\Delta)\delta_\Delta(t - k\Delta)\Delta. \tag{2.25}$$

From Figure 2.12, we see that, as in the discrete-time case [eq. (2.2)], for any value of t, only one term in the summation on the right-hand side of eq. (2.25) is nonzero.

As we let Δ approach 0, the approximation $\hat{x}(t)$ becomes better and better, and in the limit equals $x(t)$. Therefore,

$$x(t) = \lim_{\Delta \to 0} \sum_{k=-\infty}^{+\infty} x(k\Delta)\delta_\Delta(t - k\Delta)\Delta. \tag{2.26}$$

Also, as $\Delta \to 0$, the summation in eq. (2.26) approaches an integral. This can be seen by considering the graphical interpretation of the equation, illustrated in Figure 2.13. Here, we have illustrated the signals $x(\tau)$, $\delta_\Delta(t - \tau)$, and their product. We have also indicated a shaded region whose area approaches the area under $x(\tau)\delta_\Delta(t - \tau)$ as $\Delta \to 0$. Note that the shaded region has an area equal to $x(m\Delta)$ where $t - \Delta < m\Delta < t$. Furthermore, for this value of t, only the term with $k = m$ is nonzero in the summation in eq. (2.26), and thus, the right-hand side of this equation also equals $x(m\Delta)$. Consequently, it follows from eq. (2.26) and from the preceding argument that $x(t)$ equals the limit as $\Delta \to 0$ of the area under $x(\tau)\delta_\Delta(t - \tau)$. Moreover, from eq. (1.74), we know that the limit as $\Delta \to 0$ of $\delta_\Delta(t)$ is the unit impulse function $\delta(t)$. Consequently,

$$x(t) = \int_{-\infty}^{+\infty} x(\tau)\delta(t - \tau)d\tau. \tag{2.27}$$

As in discrete time, we refer to eq. (2.27) as the *sifting property* of the continuous-time impulse. We note that, for the specific example of $x(t) = u(t)$, eq. (2.27) becomes

$$u(t) = \int_{-\infty}^{+\infty} u(\tau)\delta(t - \tau)d\tau = \int_{0}^{\infty} \delta(t - \tau)d\tau, \tag{2.28}$$

since $u(\tau) = 0$ for $\tau < 0$ and $u(\tau) = 1$ for $\tau > 0$. Equation (2.28) is identical to eq. (1.75), derived in Section 1.4.2.

Once again, eq. (2.27) should be viewed as an idealization in the sense that, for Δ "small enough," the approximation of $x(t)$ in eq. (2.25) is essentially exact for any practical purpose. Equation (2.27) then simply represents an idealization of eq. (2.25) by taking Δ to be vanishingly small. Note also that we could have derived eq. (2.27) directly by using several of the basic properties of the unit impulse that we derived in Section 1.4.2.

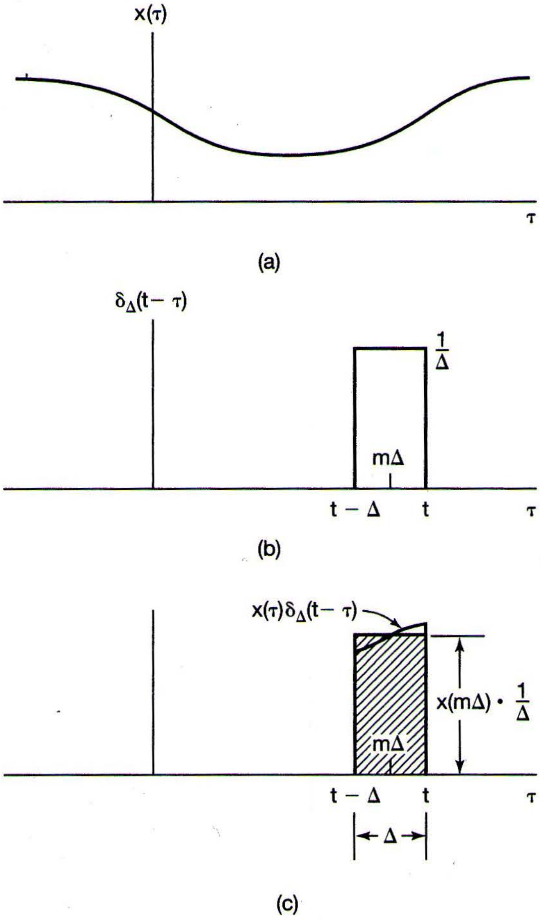


Figure 2.13 Graphical interpretation of eq. (2.26).

Specifically, as illustrated in Figure 2.14(b), the signal $\delta(t-\tau)$ (viewed as a function of τ with t fixed) is a unit impulse located at $\tau = t$. Thus, as shown in Figure 2.14(c), the signal $x(\tau)\delta(t-\tau)$ (once again viewed as a function of τ) equals $x(t)\delta(t-\tau)$ [i.e., it is a scaled impulse at $\tau = t$ with an area equal to the value of $x(t)$]. Consequently, the integral of this signal from $\tau = -\infty$ to $\tau = +\infty$ equals $x(t)$; that is,

$$\int_{-\infty}^{+\infty} x(\tau)\delta(t-\tau)d\tau = \int_{-\infty}^{+\infty} x(t)\delta(t-\tau)d\tau = x(t)\int_{-\infty}^{+\infty} \delta(t-\tau)d\tau = x(t).$$

Although this derivation follows directly from Section 1.4.2, we have included the derivation given in eqs. (2.24)–(2.27) to stress the similarities with the discrete-time case and, in particular, to emphasize the interpretation of eq. (2.27) as representing the signal $x(t)$ as a "sum" (more precisely, an integral) of weighted, shifted impulses.

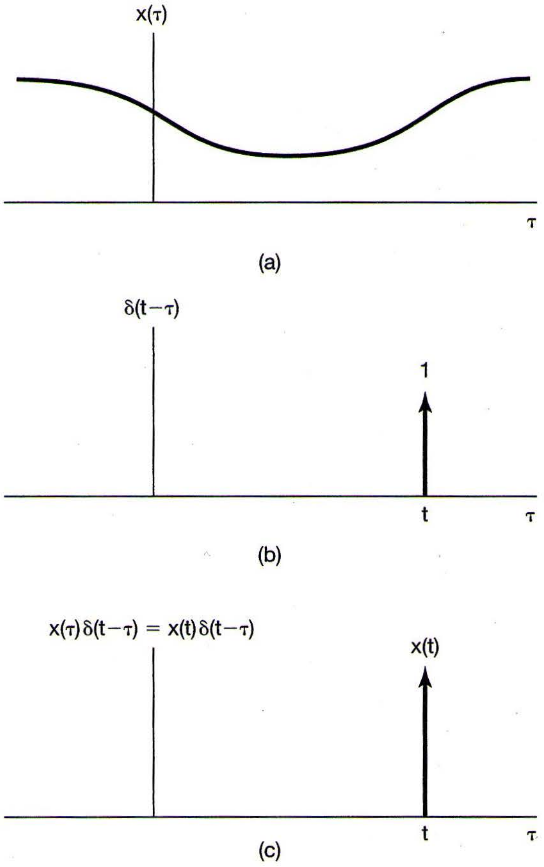


Figure 2.14 (a) Arbitrary signal $x(\tau)$; (b) impulse $\delta(t-\tau)$ as a function of τ with t fixed; (c) product of these two signals.

2.2.2 The Continuous-Time Unit Impulse Response and the Convolution Integral Representation of LTI Systems

As in the discrete-time case, the representation developed in the preceding section provides us with a way in which to view an arbitrary continuous-time signal as the superposition of scaled and shifted pulses. In particular, the approximate representation in eq. (2.25) represents the signal $\hat{x}(t)$ as a sum of scaled and shifted versions of the basic pulse signal $\delta_\Delta(t)$. Consequently, the response $\hat{y}(t)$ of a linear system to this signal will be the superposition of the responses to the scaled and shifted versions of $\delta_\Delta(t)$. Specifically, let us define $\hat{h}_{k\Delta}(t)$ as the response of an LTI system to the input $\delta_\Delta(t - k\Delta)$. Then, from eq. (2.25) and the superposition property, for continuous-time linear systems, we see that

$$\hat{y}(t) = \sum_{k=-\infty}^{+\infty} x(k\Delta)\hat{h}_{k\Delta}(t)\Delta. \tag{2.29}$$

The interpretation of eq. (2.29) is similar to that for eq. (2.3) in discrete time. In particular, consider Figure 2.15, which is the continuous-time counterpart of Figure 2.2. In

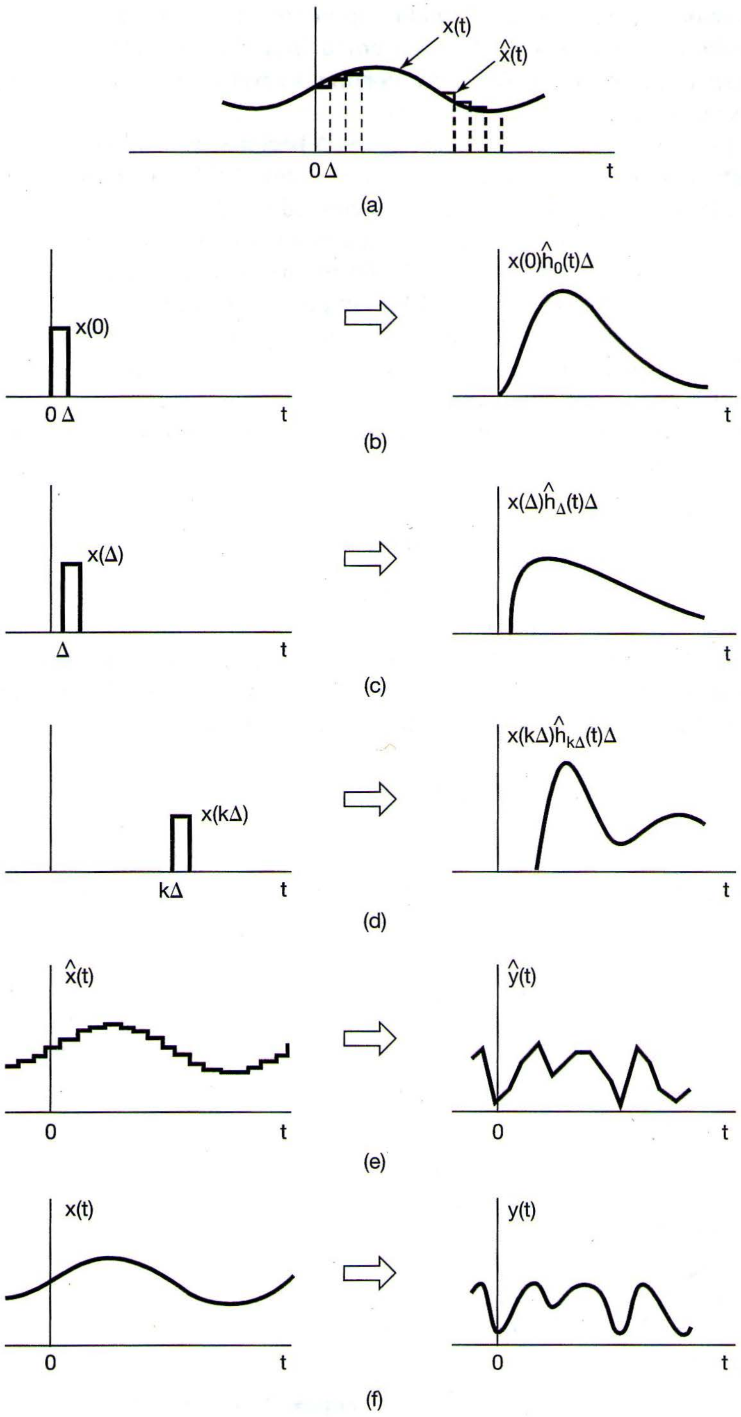

$x(t)$
$\hat{x}(t)$
0 Δ
t
(a)
$x(0)$
0 Δ
t
$x(0)\hat{h}_0(t)\Delta$
t
(b)
$x(\Delta)$
Δ
t
$x(\Delta)\hat{h}_\Delta(t)\Delta$
t
(c)
$x(k\Delta)$
kΔ
t
$x(k\Delta)\hat{h}_{k\Delta}(t)\Delta$
t
(d)
$\hat{x}(t)$
0
t
$\hat{y}(t)$
0
t
(e)
$x(t)$
0
t
$y(t)$
0
t
(f)

Figure 2.15 Graphical interpretation of the response of a continuous-time linear system as expressed in eqs. (2.29) and (2.30).

Figure 2.15(a) we have depicted the input $x(t)$ and its approximation $\hat{x}(t)$, while in Figure 2.15(b)–(d), we have shown the responses of the system to three of the weighted pulses in the expression for $\hat{x}(t)$. Then the output $\hat{y}(t)$ corresponding to $\hat{x}(t)$ is the superposition of all of these responses, as indicated in Figure 2.15(e).

What remains, then, is to consider what happens as Δ becomes vanishingly small—i.e., as $\Delta \rightarrow 0$. In particular, with $x(t)$ as expressed in eq. (2.26), $\hat{x}(t)$ becomes an increasingly good approximation to $x(t)$, and in fact, the two coincide as $\Delta \rightarrow 0$. Consequently, the response to $\hat{x}(t)$, namely, $\hat{y}(t)$ in eq. (2.29), must converge to $y(t)$, the response to the actual input $x(t)$, as illustrated in Figure 2.15(f). Furthermore, as we have said, for Δ "small enough," the duration of the pulse $\delta_\Delta(t - k\Delta)$ is of no significance, in that, as far as the system is concerned, the response to this pulse is essentially the same as the response to a unit impulse at the same point in time. That is, since the pulse $\delta_\Delta(t - k\Delta)$ corresponds to a shifted unit impulse as $\Delta \rightarrow 0$, the response $\hat{h}_{k\Delta}(t)$ to this input pulse becomes the response to an impulse in the limit. Therefore, if we let $h_\tau(t)$ denote the response at time t to a unit impulse $\delta(t - \tau)$ located at time τ, then

$$y(t) = \lim_{\Delta \to 0} \sum_{k=-\infty}^{+\infty} x(k\Delta)\hat{h}_{k\Delta}(t)\Delta. \tag{2.30}$$

As $\Delta \rightarrow 0$, the summation on the right-hand side becomes an integral, as can be seen graphically in Figure 2.16. Specifically, in Figure 2.16 the shaded rectangle represents one term in the summation on the right-hand side of eq. (2.30) and as $\Delta \rightarrow 0$ the summation approaches the area under $x(\tau)h_\tau(t)$ viewed as a function of τ. Therefore,

$$y(t) = \int_{-\infty}^{+\infty} x(\tau)h_\tau(t)d\tau. \tag{2.31}$$

The interpretation of eq. (2.31) is analogous to the one for eq. (2.29). As we showed in Section 2.2.1, any input $x(t)$ can be represented as

$$x(t) = \int_{-\infty}^{+\infty} x(\tau)\delta(t - \tau)d\tau.$$

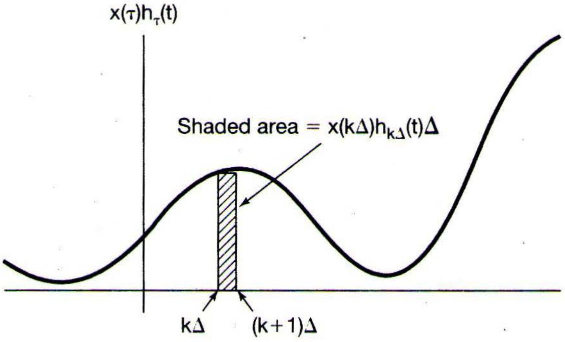


Figure 2.16 Graphical illustration of eqs. (2.30) and (2.31).

That is, we can intuitively think of $x(t)$ as a "sum" of weighted shifted impulses, where the weight on the impulse $\delta(t-\tau)$ is $x(\tau)d\tau$. With this interpretation, eq. (2.31) represents the superposition of the responses to each of these inputs, and by linearity, the weight on the response $h_\tau(t)$ to the shifted impulse $\delta(t-\tau)$ is also $x(\tau)d\tau$.

Equation (2.31) represents the general form of the response of a linear system in continuous time. If, in addition to being linear, the system is also time invariant, then $h_\tau(t) = h_0(t-\tau)$; i.e., the response of an LTI system to the unit impulse $\delta(t-\tau)$, which is shifted by τ seconds from the origin, is a similarly shifted version of the response to the unit impulse function $\delta(t)$. Again, for notational convenience, we will drop the subscript and define the *unit impulse response* $h(t)$ as

$$h(t) = h_0(t); \tag{2.32}$$

i.e., $h(t)$ is the response to $\delta(t)$. In this case, eq. (2.31) becomes

$$\boxed{y(t) = \int_{-\infty}^{+\infty} x(\tau)h(t-\tau)d\tau.} \tag{2.33}$$

Equation (2.33), referred to as the *convolution integral* or the *superposition integral*, is the continuous-time counterpart of the convolution sum of eq. (2.6) and corresponds to the representation of a continuous-time LTI system in terms of its response to a unit impulse. The convolution of two signals $x(t)$ and $h(t)$ will be represented symbolically as

$$y(t) = x(t) * h(t). \tag{2.34}$$

While we have chosen to use the same symbol $*$ to denote both discrete-time and continuous-time convolution, the context will generally be sufficient to distinguish the two cases.

As in discrete time, we see that a continuous-time LTI system is completely characterized by its impulse response—i.e., by its response to a single elementary signal, the unit impulse $\delta(t)$. In the next section, we explore the implications of this as we examine a number of the properties of convolution and of LTI systems in both continuous time and discrete time.

The procedure for evaluating the convolution integral is quite similar to that for its discrete-time counterpart, the convolution sum. Specifically, in eq. (2.33) we see that, for any value of t, the output $y(t)$ is a weighted integral of the input, where the weight on $x(\tau)$ is $h(t-\tau)$. To evaluate this integral for a specific value of t, we first obtain the signal $h(t-\tau)$ (regarded as a function of τ with t fixed) from $h(\tau)$ by a reflection about the origin and a shift to the right by t if $t > 0$ or a shift to the left by $|t|$ for $t < 0$. We next multiply together the signals $x(\tau)$ and $h(t-\tau)$, and $y(t)$ is obtained by integrating the resulting product from $\tau = -\infty$ to $\tau = +\infty$. To illustrate the evaluation of the convolution integral, let us consider several examples.

Example 2.6

Let $x(t)$ be the input to an LTI system with unit impulse response $h(t)$, where

$$x(t) = e^{-at}u(t), \quad a > 0$$

and

$$h(t) = u(t).$$

In Figure 2.17, we have depicted the functions $h(\tau)$, $x(\tau)$, and $h(t - \tau)$ for a negative value of t and for a positive value of t. From this figure, we see that for $t < 0$, the product of $x(\tau)$ and $h(t - \tau)$ is zero, and consequently, $y(t)$ is zero. For $t > 0$,

$$x(\tau)h(t - \tau) = \begin{cases} e^{-a\tau}, & 0 < \tau < t \\ 0, & \text{otherwise} \end{cases}.$$

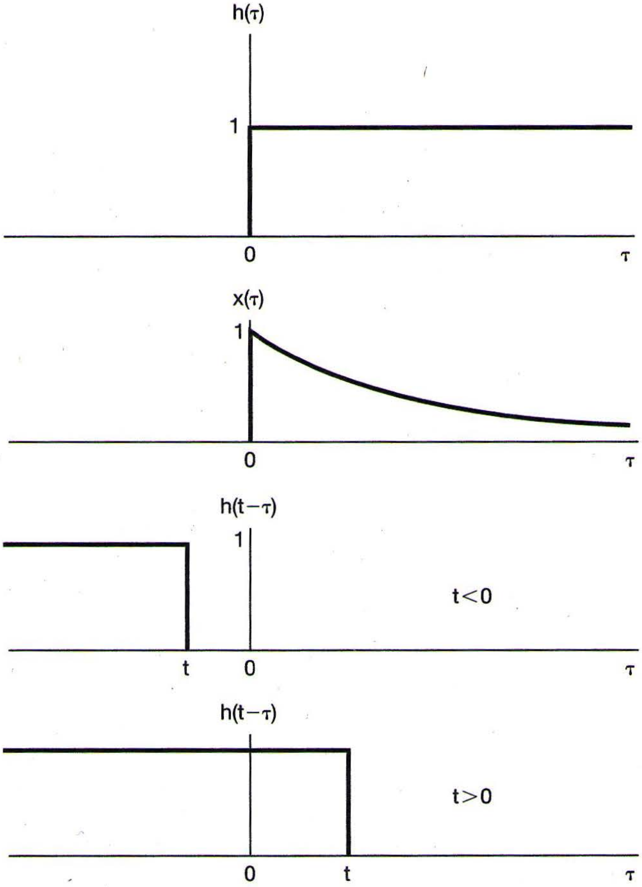


Figure 2.17 Calculation of the convolution integral for Example 2.6.

From this expression, we can compute $y(t)$ for $t > 0$:

$$y(t) = \int_0^t e^{-a\tau} d\tau = -\frac{1}{a}e^{-a\tau}\Big|_0^t$$
$$= \frac{1}{a}(1 - e^{-at}).$$

Thus, for all t, $y(t)$ is

$$y(t) = \frac{1}{a}(1 - e^{-at})u(t),$$

which is shown in Figure 2.18.

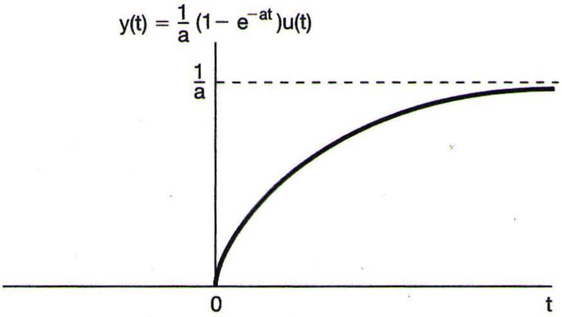


Figure 2.18 Response of the system in Example 2.6 with impulse response $h(t) = u(t)$ to the input $x(t) = e^{-at}u(t)$.

Example 2.7

Consider the convolution of the following two signals:

$$x(t) = \begin{cases} 1, & 0 < t < T \\ 0, & \text{otherwise} \end{cases},$$

$$h(t) = \begin{cases} t, & 0 < t < 2T \\ 0, & \text{otherwise} \end{cases}.$$

As in Example 2.4 for discrete-time convolution, it is convenient to consider the evaluation of $y(t)$ in separate intervals. In Figure 2.19, we have sketched $x(\tau)$ and have illustrated $h(t-\tau)$ in each of the intervals of interest. For $t < 0$ and for $t > 3T$, $x(\tau)h(t-\tau) = 0$ for all values of τ, and consequently, $y(t) = 0$. For the other intervals, the product $x(\tau)h(t-\tau)$ is as indicated in Figure 2.20. Thus, for these three intervals, the integration can be carried out graphically, with the result that

$$y(t) = \begin{cases} 0, & t < 0 \\ \frac{1}{2}t^2, & 0 < t < T \\ Tt - \frac{1}{2}T^2, & T < t < 2T \\ -\frac{1}{2}t^2 + Tt + \frac{3}{2}T^2, & 2T < t < 3T \\ 0, & 3T < t \end{cases},$$

which is depicted in Figure 2.21.

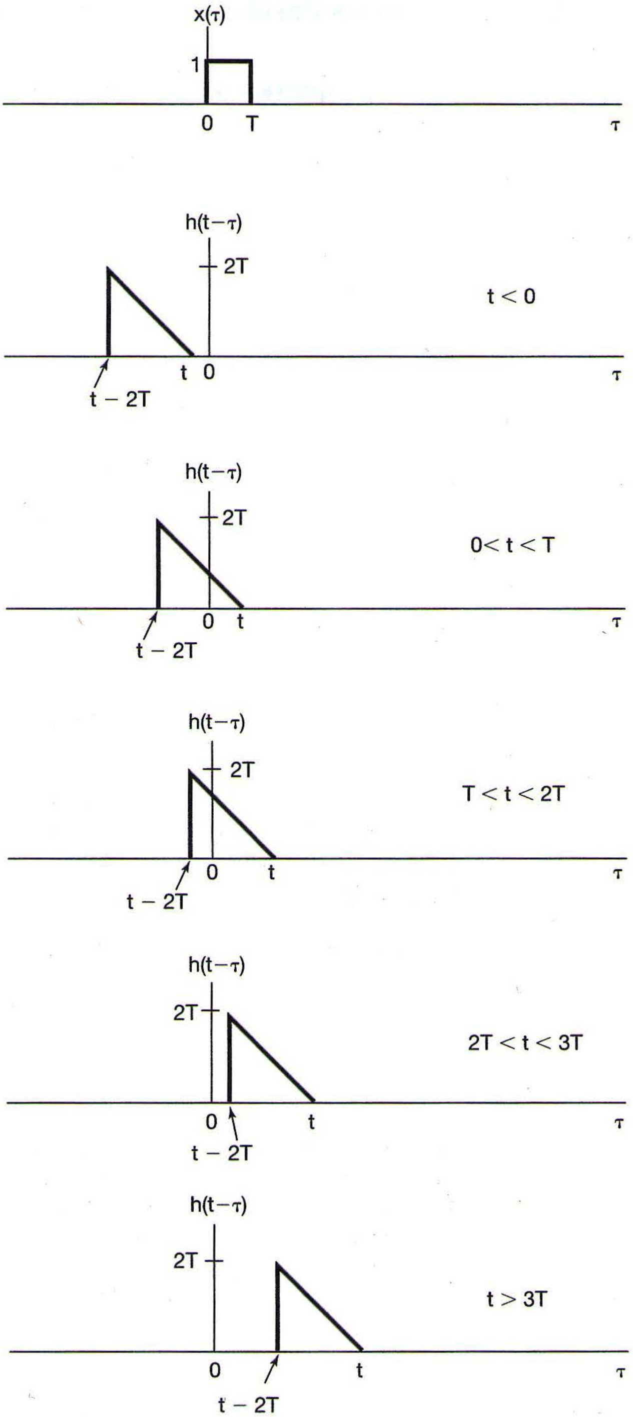

x(τ)
1
0
T
τ
h(t−τ)
2T
t < 0
t
0
t − 2T
τ
h(t−τ)
2T
0< t < T
0
t
t − 2T
τ
h(t−τ)
2T
T < t < 2T
0
t
t − 2T
τ
h(t−τ)
2T
2T < t < 3T
0
t
t − 2T
τ
h(t−τ)
2T
t > 3T
0
t
t − 2T
τ

Figure 2.19 Signals $x(\tau)$ and $h(t - \tau)$ for different values of t for Example 2.7.

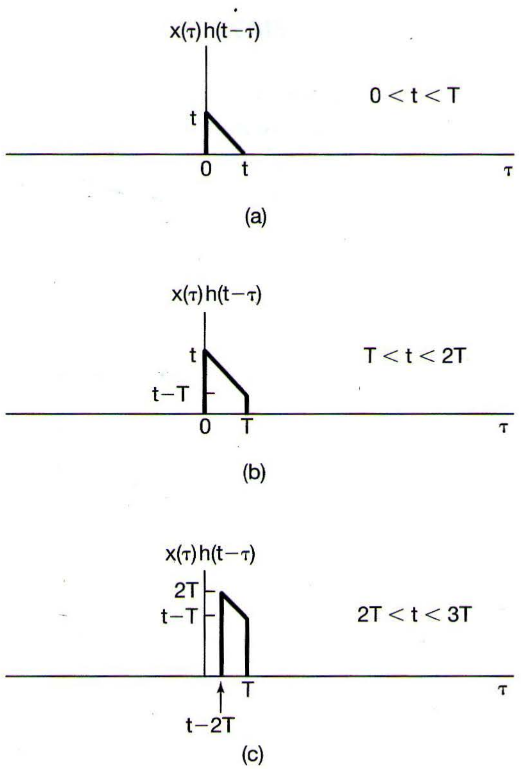

x(τ)h(t−τ)
0<t<T
t
0
t
τ
(a)
x(τ)h(t−τ)
t
t−T
T<t<2T
0
T
τ
(b)
x(τ)h(t−τ)
2T
t−T
2T<t<3T
T
τ
t−2T
(c)

Figure 2.20 Product $x(\tau)h(t-\tau)$ for Example 2.7 for the three ranges of values of t for which this product is not identically zero. (See Figure 2.19.)

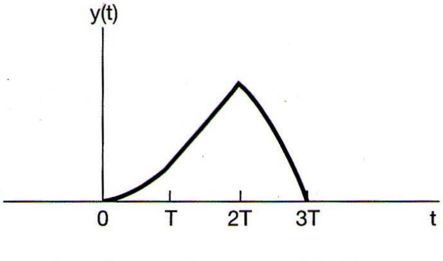


Figure 2.21 Signal $y(t) = x(t) * h(t)$ for Example 2.7.

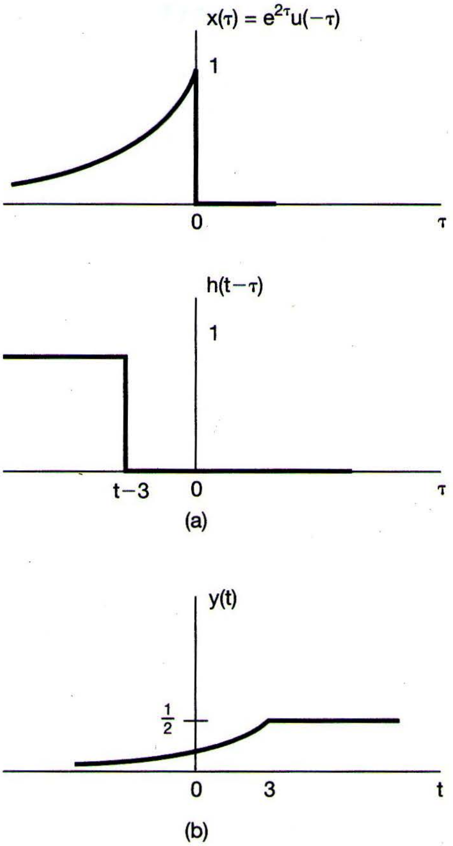
Example 2.8

Let $y(t)$ denote the convolution of the following two signals:

$$x(t) = e^{2t}u(-t), \tag{2.35}$$

$$h(t) = u(t-3). \tag{2.36}$$

The signals $x(\tau)$ and $h(t-\tau)$ are plotted as functions of τ in Figure 2.22(a). We first observe that these two signals have regions of nonzero overlap, regardless of the value



(a)

(b)

Figure 2.22 The convolution problem considered in Example 2.8.

of t. When $t - 3 \le 0$, the product of $x(\tau)$ and $h(t-\tau)$ is nonzero for $-\infty < \tau < t-3$, and the convolution integral becomes

$$y(t) = \int_{-\infty}^{t-3} e^{2\tau} d\tau = \frac{1}{2}e^{2(t-3)}. \tag{2.37}$$

For $t-3 \ge 0$, the product $x(\tau)h(t-\tau)$ is nonzero for $-\infty < \tau < 0$, so that the convolution integral is

$$y(t) = \int_{-\infty}^{0} e^{2\tau} d\tau = \frac{1}{2}. \tag{2.38}$$

The resulting signal $y(t)$ is plotted in Figure 2.22(b).

As these examples and those presented in Section 2.1 illustrate, the graphical interpretation of continuous-time and discrete-time convolution is of considerable value in visualizing the evaluation of convolution integrals and sums.

2.3 PROPERTIES OF LINEAR TIME-INVARIANT SYSTEMS

In the preceding two sections, we developed the extremely important representations of continuous-time and discrete-time LTI systems in terms of their unit impulse responses. In discrete time the representation takes the form of the convolution sum, while its continuous-time counterpart is the convolution integral, both of which we repeat here for convenience:

$$\boxed{y[n] = \sum_{k=-\infty}^{+\infty} x[k]h[n-k] = x[n] * h[n]} \tag{2.39}$$

$$\boxed{y(t) = \int_{-\infty}^{+\infty} x(\tau)h(t-\tau)d\tau = x(t) * h(t)} \tag{2.40}$$

As we have pointed out, one consequence of these representations is that the characteristics of an LTI system are completely determined by its impulse response. It is important to emphasize that this property holds in general *only* for LTI systems. In particular, as illustrated in the following example, the unit impulse response of a nonlinear system does *not* completely characterize the behavior of the system.

Example 2.9

Consider a discrete-time system with unit impulse response

$$h[n] = \begin{cases} 1, & n = 0, 1 \\ 0, & \text{otherwise} \end{cases}. \tag{2.41}$$

If the system is LTI, then eq. (2.41) completely determines its input-output behavior. In particular, by substituting eq. (2.41) into the convolution sum, eq. (2.39), we find the following explicit equation describing how the input and output of this LTI system are related:

$$y[n] = x[n] + x[n-1]. \tag{2.42}$$

On the other hand, there are *many* nonlinear systems with the same response—i.e., that given in eq. (2.41)—to the input $\delta[n]$. For example, both of the following systems have this property:

$$y[n] = (x[n] + x[n-1])^2,$$
$$y[n] = \max(x[n], x[n-1]).$$

Consequently, if the system is nonlinear it is not completely characterized by the impulse response in eq. (2.41).

The preceding example illustrates the fact that LTI systems have a number of properties not possessed by other systems, beginning with the very special representations that they have in terms of convolution sums and integrals. In the remainder of this section, we explore some of the most basic and important of these properties.

2.3.1 The Commutative Property

A basic property of convolution in both continuous and discrete time is that it is a *commutative* operation. That is, in discrete time

$$x[n] * h[n] = h[n] * x[n] = \sum_{k=-\infty}^{+\infty} h[k]x[n-k], \tag{2.43}$$

and in continuous time

$$x(t) * h(t) = h(t) * x(t) = \int_{-\infty}^{+\infty} h(\tau)x(t-\tau)d\tau. \tag{2.44}$$

These expressions can be verified in a straightforward manner by means of a substitution of variables in eqs. (2.39) and (2.40). For example, in the discrete-time case, if we let $r = n - k$ or, equivalently, $k = n - r$, eq. (2.39) becomes

$$x[n] * h[n] = \sum_{k=-\infty}^{+\infty} x[k]h[n-k] = \sum_{r=-\infty}^{+\infty} x[n-r]h[r] = h[n] * x[n]. \tag{2.45}$$

With this substitution of variables, the roles of $x[n]$ and $h[n]$ are interchanged. According to eq. (2.45), the output of an LTI system with input $x[n]$ and unit impulse response $h[n]$ is identical to the output of an LTI system with input $h[n]$ and unit impulse response $x[n]$. For example, we could have calculated the convolution in Example 2.4 by first reflecting and shifting $x[k]$, then multiplying the signals $x[n-k]$ and $h[k]$, and finally summing the products for all values of k.

Similarly, eq. (2.44) can be verified by a change of variables, and the implications of this result in continuous time are the same: The output of an LTI system with input $x(t)$ and unit impulse response $h(t)$ is identical to the output of an LTI system with input $h(t)$ and unit impulse response $x(t)$. Thus, we could have calculated the convolution in Example 2.7 by reflecting and shifting $x(t)$, multiplying the signals $x(t-\tau)$ and $h(\tau)$, and integrating over $-\infty < \tau < +\infty$. In specific cases, one of the two forms for computing convolutions [i.e., eq. (2.39) or (2.43) in discrete time and eq. (2.40) or (2.44) in continuous time] may be easier to visualize, but both forms always result in the same answer.

2.3.2 The Distributive Property

Another basic property of convolution is the *distributive* property. Specifically, convolution distributes over addition, so that in discrete time

$$x[n] * (h_1[n] + h_2[n]) = x[n] * h_1[n] + x[n] * h_2[n], \tag{2.46}$$

and in continuous time

$$x(t) * [h_1(t) + h_2(t)] = x(t) * h_1(t) + x(t) * h_2(t). \tag{2.47}$$

This property can be verified in a straightforward manner.

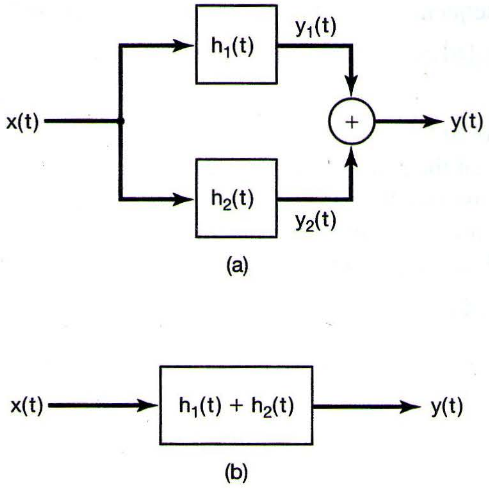


Figure 2.23 Interpretation of the distributive property of convolution for a parallel interconnection of LTI systems.

The distributive property has a useful interpretation in terms of system interconnections. Consider two continuous-time LTI systems in parallel, as indicated in Figure 2.23(a). The systems shown in the block diagram are LTI systems with the indicated unit impulse responses. This pictorial representation is a particularly convenient way in which to denote LTI systems in block diagrams, and it also reemphasizes the fact that the impulse response of an LTI system completely characterizes its behavior.

The two systems, with impulse responses $h_1(t)$ and $h_2(t)$, have identical inputs, and their outputs are added. Since

$$y_1(t) = x(t) * h_1(t)$$

and

$$y_2(t) = x(t) * h_2(t),$$

the system of Figure 2.23(a) has output

$$y(t) = x(t) * h_1(t) + x(t) * h_2(t), \tag{2.48}$$

corresponding to the right-hand side of eq. (2.47). The system of Figure 2.23(b) has output

$$y(t) = x(t) * [h_1(t) + h_2(t)], \tag{2.49}$$

corresponding to the left-hand side of eq. (2.47). Applying eq. (2.47) to eq. (2.49) and comparing the result with eq. (2.48), we see that the systems in Figures 2.23(a) and (b) are identical.

There is an identical interpretation in discrete time, in which each of the signals in Figure 2.23 is replaced by a discrete-time counterpart (i.e., $x(t)$, $h_1(t)$, $h_2(t)$, $y_1(t)$, $y_2(t)$, and $y(t)$ are replaced by $x[n]$, $h_1[n]$, $h_2[n]$, $y_1[n]$, $y_2[n]$, and $y[n]$, respectively). In summary, then, by virtue of the distributive property of convolution, a parallel combination of LTI systems can be replaced by a single LTI system whose unit impulse response is the sum of the individual unit impulse responses in the parallel combination.

Also, as a consequence of both the commutative and distributive properties, we have

$$[x_1[n] + x_2[n]] * h[n] = x_1[n] * h[n] + x_2[n] * h[n] \tag{2.50}$$

and

$$[x_1(t) + x_2(t)] * h(t) = x_1(t) * h(t) + x_2(t) * h(t), \tag{2.51}$$

which simply state that the response of an LTI system to the sum of two inputs must equal the sum of the responses to these signals individually.

As illustrated in the next example, the distributive property of convolution can also be exploited to break a complicated convolution into several simpler ones.

Example 2.10

Let $y[n]$ denote the convolution of the following two sequences:

$$x[n] = \left(\frac{1}{2}\right)^n u[n] + 2^n u[-n], \tag{2.52}$$

$$h[n] = u[n]. \tag{2.53}$$

Note that the sequence $x[n]$ is nonzero along the entire time axis. Direct evaluation of such a convolution is somewhat tedious. Instead, we may use the distributive property to express $y[n]$ as the sum of the results of two simpler convolution problems. In particular, if we let $x_1[n] = (1/2)^n u[n]$ and $x_2[n] = 2^n u[-n]$, it follows that

$$y[n] = (x_1[n] + x_2[n]) * h[n]. \tag{2.54}$$

Using the distributive property of convolution, we may rewrite eq. (2.54) as

$$y[n] = y_1[n] + y_2[n], \tag{2.55}$$

where

$$y_1[n] = x_1[n] * h[n] \tag{2.56}$$

and

$$y_2[n] = x_2[n] * h[n]. \tag{2.57}$$

The convolution in eq. (2.56) for $y_1[n]$ can be obtained from Example 2.3 (with $\alpha = 1/2$), while $y_2[n]$ was evaluated in Example 2.5. Their sum is $y[n]$, which is shown in Figure 2.24.

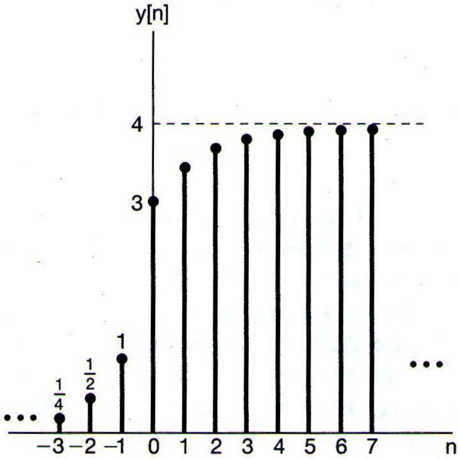


Figure 2.24 The signal $y[n] = x[n] * h[n]$ for Example 2.10.

2.3.3 The Associative Property

Another important and useful property of convolution is that it is *associative*. That is, in discrete time

$$x[n] * (h_1[n] * h_2[n]) = (x[n] * h_1[n]) * h_2[n], \tag{2.58}$$

and in continuous time

$$x(t) * [h_1(t) * h_2(t)] \doteq [x(t) * h_1(t)] * h_2(t). \tag{2.59}$$

This property is proven by straightforward manipulations of the summations and integrals involved. Examples verifying it are given in Problem 2.43.

As a consequence of the associative property, the expressions

$$y[n] = x[n] * h_1[n] * h_2[n] \tag{2.60}$$

and

$$y(t) = x(t) * h_1(t) * h_2(t) \tag{2.61}$$

are unambiguous. That is, according to eqs. (2.58) and (2.59), it does not matter in which order we convolve these signals.

An interpretation of the associative property is illustrated for discrete-time systems in Figures 2.25(a) and (b). In Figure 2.25(a),

$$\begin{aligned} y[n] &= w[n] * h_2[n] \\ &= (x[n] * h_1[n]) * h_2[n]. \end{aligned}$$

In Figure 2.25(b),

$$\begin{aligned} y[n] &= x[n] * h[n] \\ &= x[n] * (h_1[n] * h_2[n]). \end{aligned}$$

According to the associative property, the series interconnection of the two systems in Figure 2.25(a) is equivalent to the single system in Figure 2.25(b). This can be generalized to an arbitrary number of LTI systems in cascade, and the analogous interpretation and conclusion also hold in continuous time.

By using the commutative property together with the associative property, we find another very important property of LTI systems. Specifically, from Figures 2.25(a) and (b), we can conclude that the impulse response of the cascade of two LTI systems is the convolution of their individual impulse responses. Since convolution is commutative, we can compute this convolution of $h_1[n]$ and $h_2[n]$ in either order. Thus, Figures 2.25(b) and (c) are equivalent, and from the associative property, these are in turn equivalent to the system of Figure 2.25(d), which we note is a cascade combination of two systems as in Figure 2.25(a), but with the order of the cascade reversed. Consequently, the unit impulse response of a cascade of two LTI systems does not depend on the order in which they are cascaded. In fact, this holds for an arbitrary number of LTI systems in cascade: The order in which they are cascaded does not matter as far as the overall system impulse response is concerned. The same conclusions hold in continuous time as well.

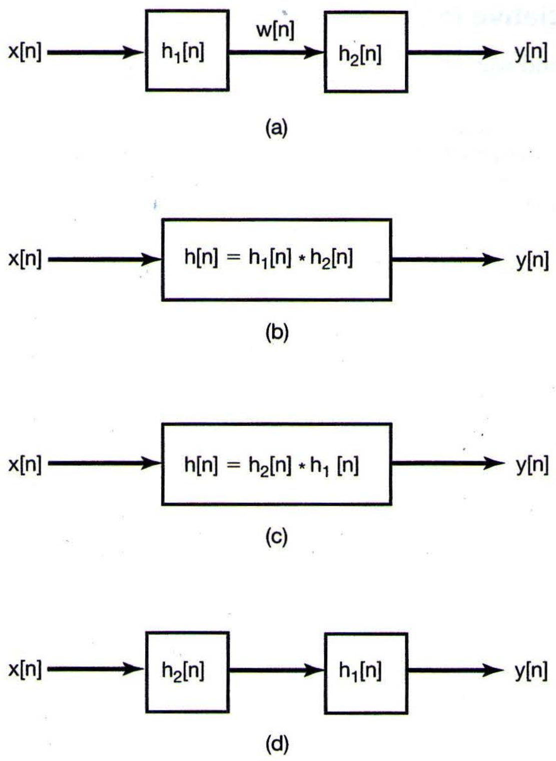


Figure 2.25 Associative property of convolution and the implication of this and the commutative property for the series interconnection of LTI systems.

It is important to emphasize that the behavior of LTI systems in cascade—and, in particular, the fact that the overall system response does not depend upon the order of the systems in the cascade—is very special to such systems. In contrast, the order in which nonlinear systems are cascaded cannot be changed, in general, without changing the overall response. For instance, if we have two memoryless systems, one being multiplication by 2 and the other squaring the input, then if we multiply first and square second, we obtain

$$y[n] = 4x^2[n].$$

However, if we multiply by 2 after squaring, we have

$$y[n] = 2x^2[n].$$

Thus, being able to interchange the order of systems in a cascade is a characteristic particular to LTI systems. In fact, as shown in Problem 2.51, we need both linearity *and* time invariance in order for this property to be true in general.

2.3.4 LTI Systems with and without Memory

As specified in Section 1.6.1, a system is memoryless if its output at any time depends only on the value of the input at that same time. From eq. (2.39), we see that the only way that this can be true for a discrete-time LTI system is if $h[n] = 0$ for $n \neq 0$. In this case the impulse response has the form

$$h[n] = K\delta[n], \tag{2.62}$$

where $K = h[0]$ is a constant, and the convolution sum reduces to the relation

$$y[n] = Kx[n]. \tag{2.63}$$

If a discrete-time LTI system has an impulse response $h[n]$ that is not identically zero for $n \neq 0$, then the system has memory. An example of an LTI system with memory is the system given by eq. (2.42). The impulse response for this system, given in eq. (2.41), is nonzero for $n = 1$.

From eq. (2.40), we can deduce similar properties for continuous-time LTI systems with and without memory. In particular, a continuous-time LTI system is memoryless if $h(t) = 0$ for $t \neq 0$, and such a memoryless LTI system has the form

$$y(t) = Kx(t) \tag{2.64}$$

for some constant K and has the impulse response

$$h(t) = K\delta(t). \tag{2.65}$$

Note that if $K = 1$ in eqs. (2.62) and (2.65), then these systems become identity systems, with output equal to the input and with unit impulse response equal to the unit impulse. In this case, the convolution sum and integral formulas imply that

$$x[n] = x[n] * \delta[n]$$

and

$$x(t) = x(t) * \delta(t),$$

which reduce to the sifting properties of the discrete-time and continuous-time unit impulses:

$$x[n] = \sum_{k=-\infty}^{+\infty} x[k]\delta[n-k]$$

$$x(t) = \int_{-\infty}^{+\infty} x(\tau)\delta(t-\tau)d\tau.$$

2.3.5 Invertibility of LTI Systems

Consider a continuous-time LTI system with impulse response $h(t)$. Based on the discussion in Section 1.6.2, this system is invertible only if an inverse system exists that, when connected in series with the original system, produces an output equal to the input to the first system. Furthermore, if an LTI system is invertible, then it has an LTI inverse. (See Problem 2.50.) Therefore, we have the picture shown in Figure 2.26. We are given a system with impulse response $h(t)$. The inverse system, with impulse response $h_1(t)$, results in $w(t) = x(t)$—such that the series interconnection in Figure 2.26(a) is identical to the

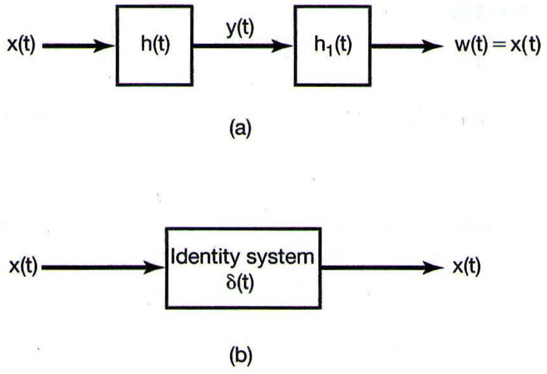


Figure 2.26 Concept of an inverse system for continuous-time LTI systems. The system with impulse response $h_1(t)$ is the inverse of the system with impulse response $h(t)$ if $h(t) * h_1(t) = \delta(t)$.

identity system in Figure 2.26(b). Since the overall impulse response in Figure 2.26(a) is $h(t) * h_1(t)$, we have the condition that $h_1(t)$ must satisfy for it to be the impulse response of the inverse system, namely,

$$h(t) * h_1(t) = \delta(t). \tag{2.66}$$

Similarly, in discrete time, the impulse response $h_1[n]$ of the inverse system for an LTI system with impulse response $h[n]$ must satisfy

$$h[n] * h_1[n] = \delta[n]. \tag{2.67}$$

The following two examples illustrate invertibility and the construction of an inverse system.

Example 2.11

Consider the LTI system consisting of a pure time shift

$$y(t) = x(t - t_0). \tag{2.68}$$

Such a system is a *delay* if $t_0 > 0$ and an *advance* if $t_0 < 0$. For example, if $t_0 > 0$, then the output at time t equals the value of the input at the earlier time $t - t_0$. If $t_0 = 0$, the system in eq. (2.68) is the identity system and thus is memoryless. For any other value of t_0, this system has memory, as it responds to the value of the input at a time other than the current time.

The impulse response for the system can be obtained from eq. (2.68) by taking the input equal to $\delta(t)$, i.e.,

$$h(t) = \delta(t - t_0). \tag{2.69}$$

Therefore,

$$x(t - t_0) = x(t) * \delta(t - t_0). \tag{2.70}$$

That is, the convolution of a signal with a shifted impulse simply shifts the signal.

To recover the input from the output, i.e., to invert the system, all that is required is to shift the output back. The system with this compensating time shift is then the inverse system. That is, if we take

$$h_1(t) = \delta(t + t_0),$$

then

$$h(t) * h_1(t) = \delta(t - t_0) * \delta(t + t_0) = \delta(t).$$

Similarly, a pure time shift in discrete time has the unit impulse response $\delta[n-n_0]$, so that convolving a signal with a shifted impulse is the same as shifting the signal. Furthermore, the inverse of the LTI system with impulse response $\delta[n - n_0]$ is the LTI system that shifts the signal in the opposite direction by the same amount—i.e., the LTI system with impulse response $\delta[n + n_0]$.

Example 2.12

Consider an LTI system with impulse response

$$h[n] = u[n]. \tag{2.71}$$

Using the convolution sum, we can calculate the response of this system to an arbitrary input:

$$y[n] = \sum_{k=-\infty}^{+\infty} x[k]u[n - k]. \tag{2.72}$$

Since $u[n - k]$ is 0 for $n - k < 0$ and 1 for $n - k \geq 0$, eq. (2.72) becomes

$$y[n] = \sum_{k=-\infty}^{n} x[k]. \tag{2.73}$$

That is, this system, which we first encountered in Section 1.6.1 [see eq. (1.92)], is a summer or accumulator that computes the running sum of all the values of the input up to the present time. As we saw in Section 1.6.2, such a system is invertible, and its inverse, as given by eq. (1.99), is

$$y[n] = x[n] - x[n - 1], \tag{2.74}$$

which is simply a *first difference* operation. Choosing $x[n] = \delta[n]$, we find that the impulse response of the inverse system is

$$h_1[n] = \delta[n] - \delta[n - 1]. \tag{2.75}$$

As a check that $h[n]$ in eq. (2.71) and $h_1[n]$ in eq. (2.75) are indeed the impulse responses of LTI systems that are inverses of each other, we can verify eq. (2.67) by direct calculation:

$$\begin{aligned} h[n] * h_1[n] &= u[n] * \{\delta[n] - \delta[n - 1]\} \\ &= u[n] * \delta[n] - u[n] * \delta[n - 1] \\ &= u[n] - u[n - 1] \\ &= \delta[n]. \end{aligned} \tag{2.76}$$

2.3.6 Causality for LTI Systems

In Section 1.6.3, we introduced the property of causality: The output of a causal system depends only on the present and past values of the input to the system. By using the convolution sum and integral, we can relate this property to a corresponding property of the impulse response of an LTI system. Specifically, in order for a discrete-time LTI system to be causal, $y[n]$ must not depend on $x[k]$ for $k > n$. From eq. (2.39), we see that for this to be true, all of the coefficients $h[n-k]$ that multiply values of $x[k]$ for $k > n$ must be zero. This then requires that the impulse response of a causal discrete-time LTI system satisfy the condition

$$h[n] = 0 \quad \text{for } n < 0. \tag{2.77}$$

According to eq. (2.77), the impulse response of a causal LTI system must be zero before the impulse occurs, which is consistent with the intuitive concept of causality. More generally, as shown in Problem 1.44, causality for a linear system is equivalent to the condition of *initial rest*; i.e., if the input to a causal system is 0 up to some point in time, then the output must also be 0 up to that time. It is important to emphasize that the equivalence of causality and the condition of initial rest applies only to linear systems. For example, as discussed in Section 1.6.6, the system $y[n] = 2x[n] + 3$ is not linear. However, it is causal and, in fact, memoryless. On the other hand, if $x[n] = 0$, $y[n] = 3 \neq 0$, so it does not satisfy the condition of initial rest.

For a causal discrete-time LTI system, the condition in eq. (2.77) implies that the convolution sum representation in eq. (2.39) becomes

$$y[n] = \sum_{k=-\infty}^{n} x[k]h[n-k], \tag{2.78}$$

and the alternative equivalent form, eq. (2.43), becomes

$$y[n] = \sum_{k=0}^{\infty} h[k]x[n-k]. \tag{2.79}$$

Similarly, a continuous-time LTI system is causal if

$$h(t) = 0 \quad \text{for } t < 0, \tag{2.80}$$

and in this case the convolution integral is given by

$$y(t) = \int_{-\infty}^{t} x(\tau)h(t-\tau)d\tau = \int_{0}^{\infty} h(\tau)x(t-\tau)d\tau. \tag{2.81}$$

Both the accumulator ($h[n] = u[n]$) and its inverse ($h[n] = \delta[n] - \delta[n-1]$), described in Example 2.12, satisfy eq. (2.77) and therefore are causal. The pure time shift with impulse response $h(t) = \delta(t-t_0)$ is causal for $t_0 \geq 0$ (when the time shift is a delay), but is noncausal for $t_0 < 0$ (in which case the time shift is an advance, so that the output anticipates future values of the input).

Finally, while causality is a property of systems, it is common terminology to refer to a signal as being causal if it is zero for $n < 0$ or $t < 0$. The motivation for this terminology comes from eqs. (2.77) and (2.80): Causality of an LTI system is equivalent to its impulse response being a causal signal.

2.3.7 Stability for LTI Systems

Recall from Section 1.6.4 that a system is *stable* if every bounded input produces a bounded output. In order to determine conditions under which LTI systems are stable, consider an input $x[n]$ that is bounded in magnitude:

$$|x[n]| < B \quad \text{for all } n. \tag{2.82}$$

Suppose that we apply this input to an LTI system with unit impulse response $h[n]$. Then, using the convolution sum, we obtain an expression for the magnitude of the output:

$$|y[n]| = \left| \sum_{k=-\infty}^{+\infty} h[k]x[n-k] \right|. \tag{2.83}$$

Since the magnitude of the sum of a set of numbers is no larger than the sum of the magnitudes of the numbers, it follows from eq. (2.83) that

$$|y[n]| \leq \sum_{k=-\infty}^{+\infty} |h[k]||x[n-k]|. \tag{2.84}$$

From eq. (2.82), $|x[n-k]| < B$ for all values of k and n. Together with eq. (2.84), this implies that

$$|y[n]| \leq B \sum_{k=-\infty}^{+\infty} |h[k]| \quad \text{for all } n. \tag{2.85}$$

From eq. (2.85), we can conclude that if the impulse response is *absolutely summable*, that is, if

$$\sum_{k=-\infty}^{+\infty} |h[k]| < \infty, \tag{2.86}$$

then $y[n]$ is bounded in magnitude, and hence, the system is stable. Therefore, eq. (2.86) is a sufficient condition to guarantee the stability of a discrete-time LTI system. In fact, this condition is also a necessary condition, since, as shown in Problem 2.49, if eq. (2.86) is not satisfied, there are bounded inputs that result in unbounded outputs. Thus, the stability of a discrete-time LTI system is completely equivalent to eq. (2.86).

In continuous time, we obtain an analogous characterization of stability in terms of the impulse response of an LTI system. Specifically, if $|x(t)| < B$ for all t, then, in analogy with eqs. (2.83)–(2.85), it follows that

$$\begin{aligned}|y(t)| &= \left|\int_{-\infty}^{+\infty} h(\tau)x(t-\tau)d\tau\right| \\ &\le \int_{-\infty}^{+\infty} |h(\tau)||x(t-\tau)|d\tau \\ &\le B\int_{-\infty}^{+\infty} |h(\tau)|d\tau.\end{aligned}$$

Therefore, the system is stable if the impulse response is *absolutely integrable*, i.e., if

$$\int_{-\infty}^{+\infty} |h(\tau)|d\tau < \infty. \tag{2.87}$$

As in discrete time, if eq. (2.87) is not satisfied, there are bounded inputs that produce unbounded outputs; therefore, the stability of a continuous-time LTI system is equivalent to eq. (2.87). The use of eqs (2.86) and (2.87) to test for stability is illustrated in the next two examples.

Example 2.13

Consider a system that is a pure time shift in either continuous time or discrete time. Then, in discrete time

$$\sum_{n=-\infty}^{+\infty} |h[n]| = \sum_{n=-\infty}^{+\infty} |\delta[n-n_0]| = 1, \tag{2.88}$$

while in continuous time

$$\int_{-\infty}^{+\infty} |h(\tau)|d\tau = \int_{-\infty}^{+\infty} |\delta(\tau - t_0)|d\tau = 1, \tag{2.89}$$

and we conclude that both of these systems are stable. This should not be surprising, since if a signal is bounded in magnitude, so is any time-shifted version of that signal.

Now consider the accumulator described in Example 2.12. As we discussed in Section 1.6.4, this is an unstable system, since, if we apply a constant input to an accumulator, the output grows without bound. That this system is unstable can also be seen from the fact that its impulse response $u[n]$ is not absolutely summable:

$$\sum_{n=-\infty}^{\infty} |u[n]| = \sum_{n=0}^{\infty} u[n] = \infty.$$

Similarly, consider the integrator, the continuous-time counterpart of the accumulator:

$$y(t) = \int_{-\infty}^{t} x(\tau)d\tau. \tag{2.90}$$

This is an unstable system for precisely the same reason as that given for the accumulator; i.e., a constant input gives rise to an output that grows without bound. The impulse response for the integrator can be found by letting $x(t) = \delta(t)$, in which case

$$h(t) = \int_{-\infty}^{t} \delta(\tau)d\tau = u(t)$$

and

$$\int_{-\infty}^{+\infty} |u(\tau)|d\tau = \int_{0}^{+\infty} d\tau = \infty.$$

Since the impulse response is not absolutely integrable, the system is not stable.

2.3.8 The Unit Step Response of an LTI System

Up to now, we have seen that the representation of an LTI system in terms of its unit impulse response allows us to obtain very explicit characterizations of system properties. Specifically, since $h[n]$ or $h(t)$ completely determines the behavior of an LTI system, we have been able to relate system properties such as stability and causality to properties of the impulse response.

There is another signal that is also used quite often in describing the behavior of LTI systems: the *unit step response*, $s[n]$ or $s(t)$, corresponding to the output when $x[n] = u[n]$ or $x(t) = u(t)$. We will find it useful on occasion to refer to the step response, and therefore, it is worthwhile relating it to the impulse response. From the convolution-sum representation, the step response of a discrete-time LTI system is the convolution of the unit step with the impulse response; that is,

$$s[n] = u[n] * h[n].$$

However, by the commutative property of convolution, $s[n] = h[n] * u[n]$, and therefore, $s[n]$ can be viewed as the response to the input $h[n]$ of a discrete-time LTI system with unit impulse response $u[n]$. As we have seen in Example 2.12, $u[n]$ is the unit impulse response of the accumulator. Therefore,

$$s[n] = \sum_{k=-\infty}^{n} h[k]. \tag{2.91}$$

From this equation and from Example 2.12, it is clear that $h[n]$ can be recovered from $s[n]$ using the relation

$$h[n] = s[n] - s[n-1]. \tag{2.92}$$

That is, the step response of a discrete-time LTI system is the running sum of its impulse response [eq. (2.91)]. Conversely, the impulse response of a discrete-time LTI system is the first difference of its step response [eq. (2.92)].

Similarly, in continuous time, the step response of an LTI system with impulse response $h(t)$ is given by $s(t) = u(t) * h(t)$, which also equals the response of an integrator [with impulse response $u(t)$] to the input $h(t)$. That is, the unit step response of a continuous-time LTI system is the running integral of its impulse response, or

$$s(t) = \int_{-\infty}^{t} h(\tau)d\tau, \tag{2.93}$$

and from eq. (2.93), the unit impulse response is the first derivative of the unit step response,[1] or

$$h(t) = \frac{ds(t)}{dt} = s'(t). \tag{2.94}$$

Therefore, in both continuous and discrete time, the unit step response can also be used to characterize an LTI system, since we can calculate the unit impulse response from it. In Problem 2.45, expressions analogous to the convolution sum and convolution integral are derived for the representations of an LTI system in terms of its unit step response.

2.4 CAUSAL LTI SYSTEMS DESCRIBED BY DIFFERENTIAL AND DIFFERENCE EQUATIONS

An extremely important class of continuous-time systems is that for which the input and output are related through a *linear constant-coefficient differential equation*. Equations of this type arise in the description of a wide variety of systems and physical phenomena. For example, as we illustrated in Chapter 1, the response of the *RC* circuit in Figure 1.1 and the motion of a vehicle subject to acceleration inputs and frictional forces, as depicted in Figure 1.2, can both be described through linear constant-coefficient differential equations. Similar differential equations arise in the description of mechanical systems containing restoring and damping forces, in the kinetics of chemical reactions, and in many other contexts as well.

Correspondingly, an important class of discrete-time systems is that for which the input and output are related through a *linear constant-coefficient difference equation*. Equations of this type are used to describe the sequential behavior of many different processes. For instance, in Example 1.10 we saw how difference equations arise in describing the accumulation of savings in a bank account, and in Example 1.11 we saw how they can be used to describe a digital simulation of a continuous-time system described by a differential equation. Difference equations also arise quite frequently in the specification of discrete-time systems designed to perform particular operations on the input signal. For example, the system that calculates the difference between successive input values, as in eq. (1.99), and the system described by eq. (1.104) that computes the average value of the input over an interval are described by difference equations.

Throughout this book, there will be many occasions in which we will consider and examine systems described by linear constant-coefficient differential and difference equations. In this section we take a first look at these systems to introduce some of the basic ideas involved in solving differential and difference equations and to uncover and explore some of the properties of systems described by such equations. In subsequent chapters, we develop additional tools for the analysis of signals and systems that will add considerably both to our ability to analyze systems described by such equations and to our understanding of their characteristics and behavior.

---

[1] Throughout this book, we will use both the notations indicated in eq. (2.94) to denote first derivatives. Analogous notation will also be used for higher derivatives.

2.4.1 Linear Constant-Coefficient Differential Equations

To introduce some of the important ideas concerning systems specified by linear constant-coefficient differential equations, let us consider a first-order differential equation as in eq. (1.85), viz.,

$$\frac{dy(t)}{dt} + 2y(t) = x(t), \tag{2.95}$$

where $y(t)$ denotes the output of the system and $x(t)$ is the input. For example, comparing eq. (2.95) to the differential equation (1.84) for the velocity of a vehicle subject to applied and frictional forces, we see that eq. (2.95) would correspond exactly to this system if $y(t)$ were identified with the vehicle's velocity $v(t)$, if $x(t)$ were taken as the applied force $f(t)$, and if the parameters in eq. (1.84) were normalized in units such that $b/m = 2$ and $1/m = 1$.

A very important point about differential equations such as eq. (2.95) is that they provide an *implicit* specification of the system. That is, they describe a relationship between the input and the output, rather than an explicit expression for the system output as a function of the input. In order to obtain an explicit expression, we must solve the differential equation. To find a solution, we need more information than that provided by the differential equation alone. For example, to determine the speed of an automobile at the end of a 10-second interval when it has been subjected to a constant acceleration of 1 m/sec^2 for 10 seconds, we would also need to know how fast the vehicle was moving at the *start* of the interval. Similarly, if we are told that a constant source voltage of 1 volt is applied to the *RC* circuit in Figure 1.1 for 10 seconds, we cannot determine what the capacitor voltage is at the end of that interval without also knowing what the initial capacitor voltage is.

More generally, to solve a differential equation, we must specify one or more auxiliary conditions, and once these are specified, we can then, in principle, obtain an explicit expression for the output in terms of the input. In other words, a differential equation such as eq. (2.95) describes a constraint between the input and the output of a system, but to characterize the system completely, we must also specify auxiliary conditions. Different choices for these auxiliary conditions then lead to different relationships between the input and the output. For the most part, in this book we will focus on the use of differential equations to describe causal LTI systems, and for such systems the auxiliary conditions take a particular, simple form. To illustrate this and to uncover some of the basic properties of the solutions to differential equations, let us take a look at the solution of eq. (2.95) for a specific input signal $x(t)$.[2]

---

[2] Our discussion of the solution of linear constant-coefficient differential equations is brief, since we assume that the reader has some familiarity with this material. For review, we recommend a text on the solution of ordinary differential equations, such as *Ordinary Differential Equations* (3rd ed.), by G. Birkhoff and G.-C. Rota (New York: John Wiley and Sons, 1978), or *Elementary Differential Equations* (3rd ed.), by W.E. Boyce and R.C. DiPrima (New York: John Wiley and Sons, 1977). There are also numerous texts that discuss differential equations in the context of circuit theory. See, for example, *Basic Circuit Theory*, by L.O. Chua, C.A. Desoer, and E.S. Kuh (New York: McGraw-Hill Book Company, 1987). As mentioned in the text, in the following chapters we present other very useful methods for solving linear differential equations that will be sufficient for our purposes. In addition, a number of exercises involving the solution of differential equations are included in the problems at the end of the chapter.

Example 2.14

Consider the solution of eq. (2.95) when the input signal is

$$x(t) = Ke^{3t}u(t), \tag{2.96}$$

where K is a real number.

The complete solution to eq. (2.96) consists of the sum of a *particular solution,* $y_p(t)$, and a *homogeneous solution,* $y_h(t)$, i.e.,

$$y(t) = y_p(t) + y_h(t), \tag{2.97}$$

where the particular solution satisfies eq. (2.95) and $y_h(t)$ is a solution of the homogeneous differential equation

$$\frac{dy(t)}{dt} + 2y(t) = 0. \tag{2.98}$$

A common method for finding the particular solution for an exponential input signal as in eq. (2.96) is to look for a so-called *forced response*—i.e., a signal of the same form as the input. With regard to eq. (2.95), since $x(t) = Ke^{3t}$ for $t > 0$, we hypothesize a solution for $t > 0$ of the form

$$y_p(t) = Ye^{3t}, \tag{2.99}$$

where Y is a number that we must determine. Substituting eqs. (2.96) and (2.99) into eq. (2.95) for $t > 0$ yields

$$3Ye^{3t} + 2Ye^{3t} = Ke^{3t}. \tag{2.100}$$

Canceling the factor e^{3t} from both sides of eq. (2.100), we obtain

$$3Y + 2Y = K, \tag{2.101}$$

or

$$Y = \frac{K}{5}, \tag{2.102}$$

so that

$$y_p(t) = \frac{K}{5}e^{3t}, \quad t > 0. \tag{2.103}$$

In order to determine $y_h(t)$, we hypothesize a solution of the form

$$y_h(t) = Ae^{st}. \tag{2.104}$$

Substituting this into eq. (2.98) gives

$$Ase^{st} + 2Ae^{st} = Ae^{st}(s + 2) = 0. \tag{2.105}$$

From this equation, we see that we must take $s = -2$ and that Ae^{-2t} is a solution to eq. (2.98) for *any* choice of A. Utilizing this fact and eq. (2.103) in eq. (2.97), we find that the solution of the differential equation for $t > 0$ is

$$y(t) = Ae^{-2t} + \frac{K}{5}e^{3t}, \quad t > 0. \tag{2.106}$$

As noted earlier, the differential equation (2.95) by itself does not specify uniquely the response $y(t)$ to the input $x(t)$ in eq. (2.96). In particular, the constant A in eq. (2.106) has not yet been determined. In order for the value of A to be determined, we need to specify an auxiliary condition in addition to the differential equation (2.95). As explored in Problem 2.34, different choices for this auxiliary condition lead to different solutions $y(t)$ and, consequently, to different relationships between the input and the output. As we have indicated, for the most part in this book we focus on differential and difference equations used to describe systems that are LTI and causal, and in this case the auxiliary condition takes the form of the condition of initial rest. That is, as shown in Problem 1.44, for a causal LTI system, if $x(t) = 0$ for $t < t_0$, then $y(t)$ must also equal 0 for $t < t_0$. From eq. (2.96), we see that for our example $x(t) = 0$ for $t < 0$, and thus, the condition of initial rest implies that $y(t) = 0$ for $t < 0$. Evaluating eq. (2.106) at $t = 0$ and setting $y(0) = 0$ yields

$$0 = A + \frac{K}{5},$$

or

$$A = -\frac{K}{5}.$$

Thus, for $t > 0$,

$$y(t) = \frac{K}{5}\left[e^{3t} - e^{-2t}\right], \tag{2.107}$$

while for $t < 0$, $y(t) = 0$, because of the condition of initial rest. Combining these two cases, we obtain the full solution

$$y(t) = \frac{K}{5}\left[e^{3t} - e^{-2t}\right]u(t). \tag{2.108}$$

Example 2.14 illustrates several very important points concerning linear constant-coefficient differential equations and the systems they represent. First, the response to an input $x(t)$ will generally consist of the sum of a particular solution to the differential equation and a homogeneous solution—i.e., a solution to the differential equation with the input set to zero. The homogeneous solution is often referred to as the *natural response* of the system. The natural responses of simple electrical circuits and mechanical systems are explored in Problems 2.61 and 2.62.

In Example 2.14 we also saw that, in order to determine completely the relationship between the input and the output of a system described by a differential equation such as eq. (2.95), we must specify auxiliary conditions. An implication of this fact, which is illustrated in Problem 2.34, is that different choices of auxiliary conditions lead to different relationships between the input and the output. As we illustrated in the example, for the most part we will use the condition of initial rest for systems described by differential equations. In the example, since the input was 0 for $t < 0$, the condition of initial rest implied the initial condition $y(0) = 0$. As we have stated, and as illustrated in

Problem 2.33, under the condition of initial rest the system described by eq. (2.95) is LTI and causal.[3] For example, if we multiply the input in eq. (2.96) by 2, the resulting output would be twice the output in eq. (2.108).

It is important to emphasize that the condition of initial rest does not specify a zero initial condition at a fixed point in time, but rather adjusts this point in time so that the response is zero *until* the input becomes nonzero. Thus, if $x(t) = 0$ for $t \leq t_0$ for the causal LTI system described by eq. (2.95), then $y(t) = 0$ for $t \leq t_0$, and we would use the initial condition $y(t_0) = 0$ to solve for the output for $t > t_0$. As a physical example, consider again the circuit in Figure 1.1, also discussed in Example 1.8. Initial rest for this example corresponds to the statement that, until we connect a nonzero voltage source to the circuit, the capacitor voltage is zero. Thus, if we begin to use the circuit at noon today, the initial capacitor voltage as we connect the voltage source at noon today is zero. Similarly, if we begin to use the circuit at noon tomorrow instead, the initial capacitor voltage as we connect the voltage source at noon tomorrow is zero.

This example also provides us with some intuition as to why the condition of initial rest makes a system described by a linear constant-coefficient differential equation time invariant. For example, if we perform an experiment on the circuit, starting from initial rest, then, assuming that the coefficients R and C don't change over time, we would expect to get the same results whether we ran the experiment today or tomorrow. That is, if we perform identical experiments on the two days, where the circuit starts from initial rest at noon on each day, then we would expect to see identical responses—i.e., responses that are simply time-shifted by one day with respect to each other.

While we have used the first-order differential equation (2.95) as the vehicle for the discussion of these issues, the same ideas extend directly to systems described by higher order differential equations. A general Nth-order linear constant-coefficient differential equation is given by

$$\sum_{k=0}^{N} a_k \frac{d^k y(t)}{dt^k} = \sum_{k=0}^{M} b_k \frac{d^k x(t)}{dt^k}. \tag{2.109}$$

The order refers to the highest derivative of the output $y(t)$ appearing in the equation. In the case when $N = 0$, eq. (2.109) reduces to

$$y(t) = \frac{1}{a_0} \sum_{k=0}^{M} b_k \frac{d^k x(t)}{dt^k}. \tag{2.110}$$

In this case, $y(t)$ is an explicit function of the input $x(t)$ and its derivatives. For $N \geq 1$, eq. (2.109) specifies the output implicitly in terms of the input. In this case, the analysis of the equation proceeds just as in our discussion of the first-order differential equation in Example 2.14. The solution $y(t)$ consists of two parts—a particular solution to eq. (2.109) plus a solution to the homogeneous differential equation

---

[3] In fact, as is also shown in Problem 2.34, if the initial condition for eq. (2.95) is nonzero, the resulting system is incrementally linear. That is, the overall response can be viewed, much as in Figure 1.48, as the superposition of the response to the initial conditions alone (with input set to 0) and the response to the input with an initial condition of 0 (i.e., the response of the causal LTI system described by eq. (2.95)).

$$\sum_{k=0}^{N} a_k \frac{d^k y(t)}{dt^k} = 0. \tag{2.111}$$

The solutions to this equation are referred to as the *natural responses* of the system.

As in the first-order case, the differential equation (2.109) does not completely specify the output in terms of the input, and we need to identify auxiliary conditions to determine completely the input-output relationship for the system. Once again, different choices for these auxiliary conditions result in different input-output relationships, but for the most part, in this book we will use the condition of initial rest when dealing with systems described by differential equations. That is, if $x(t) = 0$ for $t \leq t_0$, we assume that $y(t) = 0$ for $t \leq t_0$, and therefore, the response for $t > t_0$ can be calculated from the differential equation (2.109) with the initial conditions

$$y(t_0) = \frac{dy(t_0)}{dt} = \ldots = \frac{d^{N-1}y(t_0)}{dt^{N-1}} = 0. \tag{2.112}$$

Under the condition of initial rest, the system described by eq. (2.109) is causal and LTI. Given the initial conditions in eq. (2.112), the output $y(t)$ can, in principle, be determined by solving the differential equation in the manner used in Example 2.14 and further illustrated in several problems at the end of the chapter. However, in Chapters 4 and 9 we will develop some tools for the analysis of continuous-time LTI systems that greatly facilitate the solution of differential equations and, in particular, provide us with powerful methods for analyzing and characterizing the properties of systems described by such equations.

2.4.2 Linear Constant-Coefficient Difference Equations

The discrete-time counterpart of eq. (2.109) is the Nth-order linear constant-coefficient difference equation

$$\sum_{k=0}^{N} a_k y[n-k] = \sum_{k=0}^{M} b_k x[n-k]. \tag{2.113}$$

An equation of this type can be solved in a manner exactly analogous to that for differential equations. (See Problem 2.32.)[4] Specifically, the solution $y[n]$ can be written as the sum of a particular solution to eq. (2.113) and a solution to the homogeneous equation

$$\sum_{k=0}^{N} a_k y[n-k] = 0. \tag{2.114}$$

[4] For a detailed treatment of the methods for solving linear constant-coefficient difference equations, see *Finite Difference Equations,* by H. Levy and F. Lessman (New York: Macmillan, Inc., 1961), or *Finite Difference Equations and Simulations* (Englewood Cliffs, NJ: Prentice-Hall, 1968) by F. B. Hildebrand. In Chapter 6, we present another method for solving difference equations that greatly facilitates the analysis of linear time-invariant systems that are so described. In addition, we refer the reader to the problems at the end of this chapter that deal with the solution of difference equations.

The solutions to this homogeneous equation are often referred to as the natural responses of the system described by eq. (2.113).

As in the continuous-time case, eq. (2.113) does not completely specify the output in terms of the input. To do this, we must also specify some auxiliary conditions. While there are many possible choices for auxiliary conditions, leading to different input-output relationships, we will focus for the most part on the condition of initial rest—i.e., if $x[n] = 0$ for $n < n_0$, then $y[n] = 0$ for $n < n_0$ as well. With initial rest, the system described by eq. (2.113) is LTI and causal.

Although all of these properties can be developed following an approach that directly parallels our discussion for differential equations, the discrete-time case offers an alternative path. This stems from the observation that eq. (2.113) can be rearranged in the form

$$y[n] = \frac{1}{a_0}\left\{\sum_{k=0}^{M} b_k x[n-k] - \sum_{k=1}^{N} a_k y[n-k]\right\}. \tag{2.115}$$

Equation (2.115) directly expresses the output at time n in terms of previous values of the input and output. From this, we can immediately see the need for auxiliary conditions. In order to calculate $y[n]$, we need to know $y[n-1], \ldots, y[n-N]$. Therefore, if we are given the input for all n and a set of auxiliary conditions such as $y[-N], y[-N+1], \ldots, y[-1]$, eq. (2.115) can be solved for successive values of $y[n]$.

An equation of the form of eq. (2.113) or eq. (2.115) is called a *recursive equation,* since it specifies a recursive procedure for determining the output in terms of the input and previous outputs. In the special case when $N = 0$, eq. (2.115) reduces to

$$y[n] = \sum_{k=0}^{M}\left(\frac{b_k}{a_0}\right)x[n-k]. \tag{2.116}$$

This is the discrete-time counterpart of the continuous-time system given in eq. (2.110). Here, $y[n]$ is an explicit function of the present and previous values of the input. For this reason, eq. (2.116) is often called a *nonrecursive equation,* since we do not recursively use previously computed values of the output to compute the present value of the output. Therefore, just as in the case of the system given in eq. (2.110), we do not need auxiliary conditions in order to determine $y[n]$. Furthermore, eq. (2.116) describes an LTI system, and by direct computation, the impulse response of this system is found to be

$$h[n] = \begin{cases} \frac{b_n}{a_0}, & 0 \leq n \leq M \\ 0, & \text{otherwise} \end{cases}. \tag{2.117}$$

That is, eq. (2.116) is nothing more than the convolution sum. Note that the impulse response for it has finite duration; that is, it is nonzero only over a finite time interval. Because of this property, the system specified by eq. (2.116) is often called a *finite impulse response (FIR) system.*

Although we do not require auxiliary conditions for the case of $N = 0$, such conditions are needed for the recursive case when $N \geq 1$. To illustrate the solution of such an equation, and to gain some insight into the behavior and properties of recursive difference equations, let us examine the following simple example:

Example 2.15

Consider the difference equation

$$y[n] - \frac{1}{2}y[n-1] = x[n]. \tag{2.118}$$

Eq. (2.118) can also be expressed in the form

$$y[n] = x[n] + \frac{1}{2}y[n-1], \tag{2.119}$$

highlighting the fact that we need the previous value of the output, $y[n-1]$, to calculate the current value. Thus, to begin the recursion, we need an initial condition.

For example, suppose that we impose the condition of initial rest and consider the input

$$x[n] = K\delta[n]. \tag{2.120}$$

In this case, since $x[n] = 0$ for $n \leq -1$, the condition of initial rest implies that $y[n] = 0$ for $n \leq -1$, so that we have as an initial condition $y[-1] = 0$. Starting from this initial condition, we can solve for successive values of $y[n]$ for $n \geq 0$ as follows:

$$y[0] = x[0] + \frac{1}{2}y[-1] = K, \tag{2.121}$$

$$y[1] = x[1] + \frac{1}{2}y[0] = \frac{1}{2}K, \tag{2.122}$$

$$y[2] = x[2] + \frac{1}{2}y[1] = \left(\frac{1}{2}\right)^2 K, \tag{2.123}$$

$$\vdots$$

$$y[n] = x[n] + \frac{1}{2}y[n-1] = \left(\frac{1}{2}\right)^n K. \tag{2.124}$$

Since the system specified by eq. (2.118) and the condition of initial rest is LTI, its input-output behavior is completely characterized by its impulse response. Setting $K = 1$, we see that the impulse response for the system considered in this example is

$$h[n] = \left(\frac{1}{2}\right)^n u[n]. \tag{2.125}$$

Note that the causal LTI system in Example 2.15 has an impulse response of infinite duration. In fact, if $N \geq 1$ in eq. (2.113), so that the difference equation is recursive, it is usually the case that the LTI system corresponding to this equation together with the condition of initial rest will have an impulse response of infinite duration. Such systems are commonly referred to as *infinite impulse response (IIR) systems.*

As we have indicated, for the most part we will use recursive difference equations in the context of describing and analyzing systems that are linear, time-invariant, and causal, and consequently, we will usually make the assumption of initial rest. In Chapters 5 and 10 we will develop tools for the analysis of discrete-time systems that will provide us with very useful and efficient methods for solving linear constant-coefficient difference equations and for analyzing the properties of the systems that they describe.

2.4.3 Block Diagram Representations of First-Order Systems Described by Differential and Difference Equations

An important property of systems described by linear constant-coefficient difference and differential equations is that they can be represented in very simple and natural ways in terms of block diagram interconnections of elementary operations. This is significant for a number of reasons. One is that it provides a pictorial representation which can add to our understanding of the behavior and properties of these systems. In addition, such representations can be of considerable value for the simulation or implementation of the systems. For example, the block diagram representation to be introduced in this section for continuous-time systems is the basis for early analog computer simulations of systems described by differential equations, and it can also be directly translated into a program for the simulation of such a system on a digital computer. In addition, the corresponding representation for discrete-time difference equations suggests simple and efficient ways in which the systems that the equations describe can be implemented in digital hardware. In this section, we illustrate the basic ideas behind these block diagram representations by constructing them for the causal first-order systems introduced in Examples 1.8–1.11. In Problems 2.57–2.60 and Chapters 9 and 10, we consider block diagrams for systems described by other, more complex differential and difference equations.

We begin with the discrete-time case and, in particular, the causal system described by the first-order difference equation

$$y[n] + ay[n-1] = bx[n]. \tag{2.126}$$

To develop a block diagram representation of this system, note that the evaluation of eq. (2.126) requires three basic operations: addition, multiplication by a coefficient, and delay (to capture the relationship between $y[n]$ and $y[n-1]$). Thus, let us define three basic network elements, as indicated in Figure 2.27. To see how these basic elements can be used to represent the causal system described by eq. (2.126), we rewrite this equation in the form that directly suggests a recursive algorithm for computing successive values of the output $y[n]$:

$$y[n] = -ay[n-1] + bx[n]. \tag{2.127}$$

This algorithm is represented pictorially in Figure 2.28, which is an example of a feedback system, since the output is fed back through a delay and a multiplication by a coefficient and is then added to $bx[n]$. The presence of feedback is a direct consequence of the recursive nature of eq. (2.127).

The block diagram in Figure 2.28 makes clear the required memory in this system and the consequent need for initial conditions. In particular, a delay corresponds to a memory element, as the element must retain the previous value of its input. Thus, the initial value of this memory element serves as a necessary initial condition for the recursive calculation specified pictorially in Figure 2.28 and mathematically in eq. (2.127). Of course, if the system described by eq. (2.126) is initially at rest, the initial value stored in the memory element is zero.

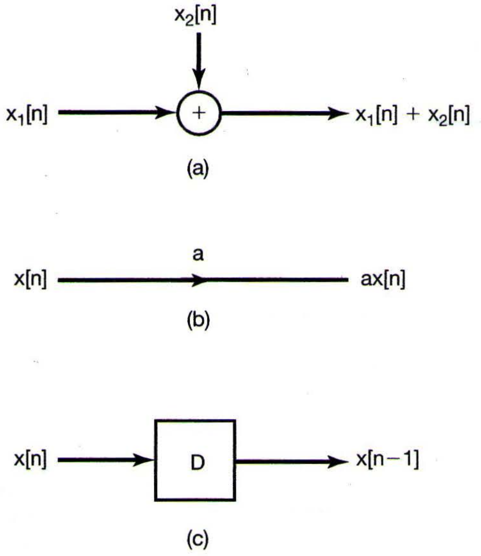


Figure 2.27 Basic elements for the block diagram representation of the causal system described by eq. (2.126): (a) an adder; (b) multiplication by a coefficient; (c) a unit delay.

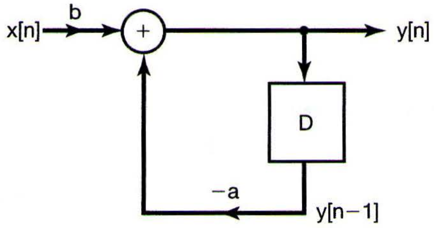


Figure 2.28 Block diagram representation for the causal discrete-time system described by eq. (2.126).

Consider next the causal continuous-time system described by a first-order differential equation:

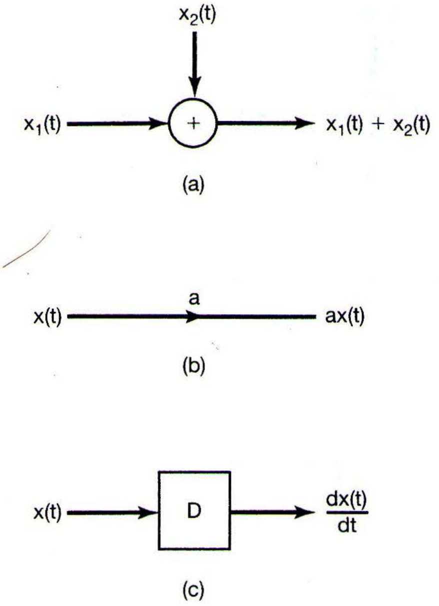
$$\frac{dy(t)}{dt} + ay(t) = bx(t). \tag{2.128}$$

As a first attempt at defining a block diagram representation for this system, let us rewrite it as

$$y(t) = -\frac{1}{a}\frac{dy(t)}{dt} + \frac{b}{a}x(t). \tag{2.129}$$

The right-hand side of this equation involves three basic operations: addition, multiplication by a coefficient, and differentiation. Therefore, if we define the three basic network elements indicated in Figure 2.29, we can consider representing eq. (2.129) as an interconnection of these basic elements in a manner analogous to that used for the discrete-time system described previously, resulting in the block diagram of Figure 2.30.

While the latter figure is a valid representation of the causal system described by eq. (2.128), it is not the representation that is most frequently used or the representation that leads directly to practical implementations, since differentiators are both difficult to implement and extremely sensitive to errors and noise. An alternative implementation that



(a)

(b)

(c)

Figure 2.29 One possible set of basic elements for the block diagram representation of the continuous-time system described by eq. (2.128): (a) an adder; (b) multiplication by a coefficient; (c) a differentiator.

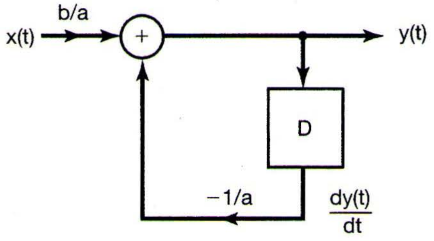


Figure 2.30 Block diagram representation for the system in eqs. (2.128) and (2.129), using adders, multiplications by coefficients, and differentiators.

is much more widely used can be obtained by first rewriting eq. (2.128) as

$$\frac{dy(t)}{dt} = bx(t) - ay(t) \tag{2.130}$$

and then integrating from $-\infty$ to t. Specifically, if we assume that in the system described by eq. (2.130) the value of $y(-\infty)$ is zero, then the integral of $dy(t)/dt$ from $-\infty$ to t is precisely $y(t)$. Consequently, we obtain the equation

$$y(t) = \int_{-\infty}^{t} \left[bx(\tau) - ay(\tau)\right] d\tau. \tag{2.131}$$

In this form, our system can be implemented using the adder and coefficient multiplier indicated in Figure 2.29, together with an *integrator*, as defined in Figure 2.31. Figure 2.32 is a block diagram representation for this system using these elements.

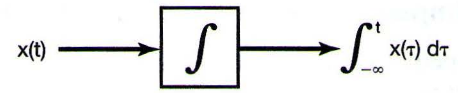


Figure 2.31 Pictorial representation of an integrator.

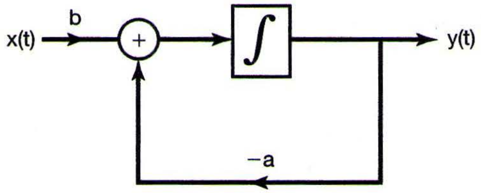


Figure 2.32 Block diagram representation for the system in eqs. (2.128) and (2.131), using adders, multiplications by coefficients, and integrators.

Since integrators can be readily implemented using operational amplifiers, representations such as that in Figure 2.32 lead directly to analog implementations, and indeed, this is the basis for both early analog computers and modern analog computation systems. Note that in the continuous-time case it is the integrator that represents the memory storage element of the system. This is perhaps more readily seen if we consider integrating eq. (2.130) from a finite point in time t_0, resulting in the expression

$$y(t) = y(t_0) + \int_{t_0}^{t} \left[bx(\tau) - ay(\tau)\right] d\tau. \tag{2.132}$$

Equation (2.132) makes clear the fact that the specification of $y(t)$ requires an initial condition, namely, the value of $y(t_0)$. It is precisely this value that the integrator stores at time t_0.

While we have illustrated block diagram constructions only for the simplest first-order differential and difference equations, such block diagrams can also be developed for higher order systems, providing both valuable intuition for and possible implementations of these systems. Examples of block diagrams for higher order systems can be found in Problems 2.58 and 2.60.

2.5 SINGULARITY FUNCTIONS

In this section, we take another look at the continuous-time unit impulse function in order to gain additional intuitions about this important idealized signal and to introduce a set of related signals known collectively as *singularity functions*. In particular, in Section 1.4.2 we suggested that a continuous-time unit impulse could be viewed as the idealization of a pulse that is "short enough" so that its shape and duration is of no practical consequence—i.e., so that as far as the response of any particular LTI system is concerned, all of the area under the pulse can be thought of as having been applied instantaneously. In this section, we would first like to provide a concrete example of what this means and then use the interpretation embodied within the example to show that the key to the use of unit impulses and other singularity functions is in the specification of how LTI systems respond to these idealized signals; i.e., the signals are in essence defined in terms of how they behave under convolution with other signals.

2.5.1 The Unit Impulse as an Idealized Short Pulse

From the sifting property, eq. (2.27), the unit impulse $\delta(t)$ is the impulse response of the identity system. That is,

$$x(t) = x(t) * \delta(t) \tag{2.133}$$

for any signal $x(t)$. Therefore, if we take $x(t) = \delta(t)$, we have

$$\delta(t) = \delta(t) * \delta(t). \tag{2.134}$$

Equation (2.134) is a basic property of the unit impulse, and it also has a significant implication for our interpretation of the unit impulse as an idealized pulse. For example, as in Section 1.4.2, suppose that we think of $\delta(t)$ as the limiting form of a rectangular pulse. Specifically, let $\delta_\Delta(t)$ correspond to the rectangular pulse defined in Figure 1.34, and let

$$r_\Delta(t) = \delta_\Delta(t) * \delta_\Delta(t). \tag{2.135}$$

Then $r_\Delta(t)$ is as sketched in Figure 2.33. If we wish to interpret $\delta(t)$ as the limit as $\Delta \rightarrow 0$ of $\delta_\Delta(t)$, then, by virtue of eq. (2.134), the limit as $\Delta \rightarrow 0$ for $r_\Delta(t)$ must also be a unit impulse. In a similar manner, we can argue that the limits as $\Delta \rightarrow 0$ of $r_\Delta(t) * r_\Delta(t)$ or $r_\Delta(t) * \delta_\Delta(t)$ must be unit impulses, and so on. Thus, we see that for consistency, if we define the unit impulse as the limiting form of some signal, then in fact, there is an unlimited number of very dissimilar-looking signals, all of which behave like an impulse in the limit.

The key words in the preceding paragraph are "behave like an impulse," where, as we have indicated, what we mean by this is that the response of an LTI system to all of these signals is essentially identical, as long as the pulse is "short enough," i.e., Δ is "small enough." The following example illustrates this idea:

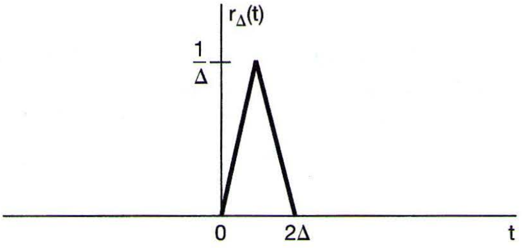


Figure 2.33 The signal $r_\Delta(t)$ defined in eq. (2.135).

Example 2.16

Consider the LTI system described by the first-order differential equation

$$\frac{dy(t)}{dt} + 2y(t) = x(t), \tag{2.136}$$

together with the condition of initial rest. Figure 2.34 depicts the response of this system to $\delta_\Delta(t)$, $r_\Delta(t)$, $r_\Delta(t) * \delta_\Delta(t)$, and $r_\Delta(t) * r_\Delta(t)$ for several values of Δ. For Δ large enough, the responses to these input signals differ noticeably. However, for Δ sufficiently small, the responses are essentially indistinguishable, so that all of the input signals "behave" in the same way. Furthermore, as suggested by the figure, the limiting form of all of these responses is precisely $e^{-2t}u(t)$. Since the limit of each of these signals as $\Delta \rightarrow 0$ is the unit impulse, we conclude that $e^{-2t}u(t)$ is the impulse response for this system.[5]

---

[5] In Chapters 4 and 9, we will describe much simpler ways to determine the impulse response of causal LTI systems described by linear constant-coefficient differential equations.

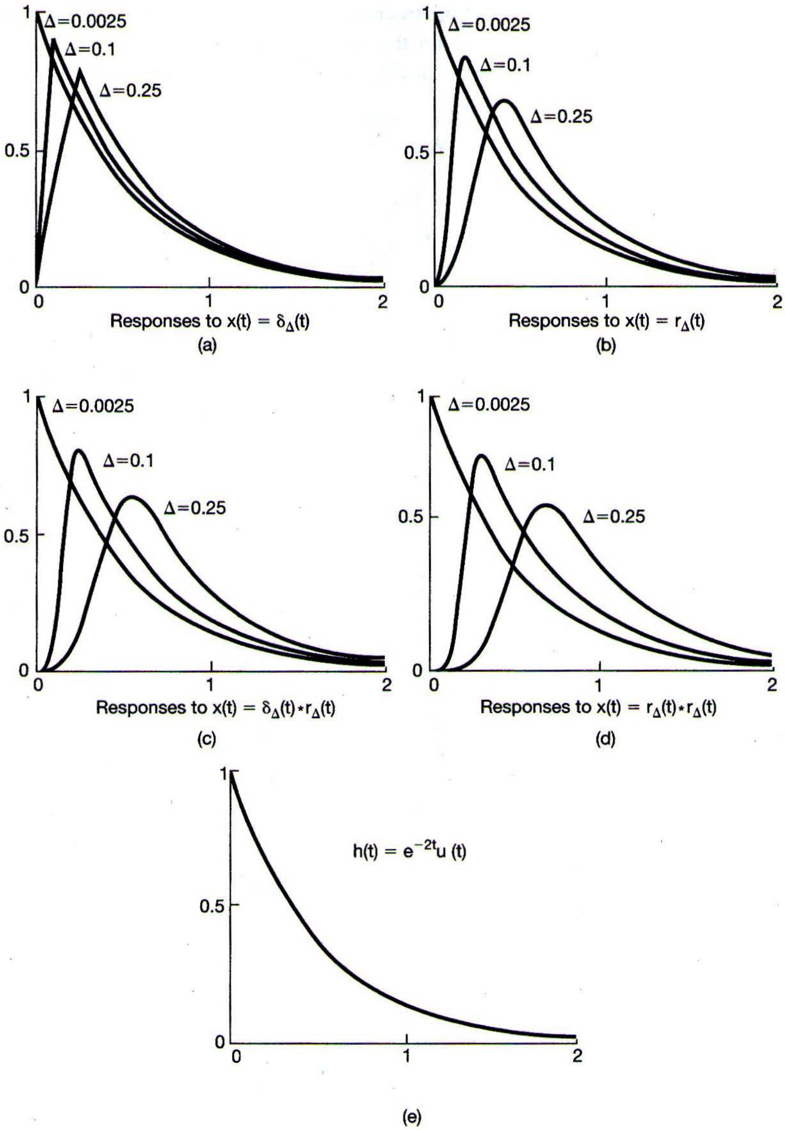


Figure 2.34 Interpretation of a unit impulse as the idealization of a pulse whose duration is "short enough" so that, as far as the response of an LTI system to this pulse is concerned, the pulse can be thought of as having been applied instantaneously: (a) responses of the causal LTI system described by eq. (2.136) to the input $\delta_\Delta(t)$ for $\Delta = 0.25$, 0.1, and 0.0025; (b) responses of the same system to $r_\Delta(t)$ for the same values of Δ; (c) responses to $\delta_\Delta(t) * r_\Delta(t)$; (d) responses to $r_\Delta(t) * r_\Delta(t)$; (e) the impulse response $h(t) = e^{-2t}u(t)$ for the system. Note that, for $\Delta = 0.25$, there are noticeable differences among the responses to these different signals; however, as Δ becomes smaller, the differences diminish, and all of the responses converge to the impulse response shown in (e).

One important point to be emphasized is that what we mean by "Δ small enough" depends on the particular LTI system to which the preceding pulses are applied. For example, in Figure 2.35, we have illustrated the responses to these pulses for different

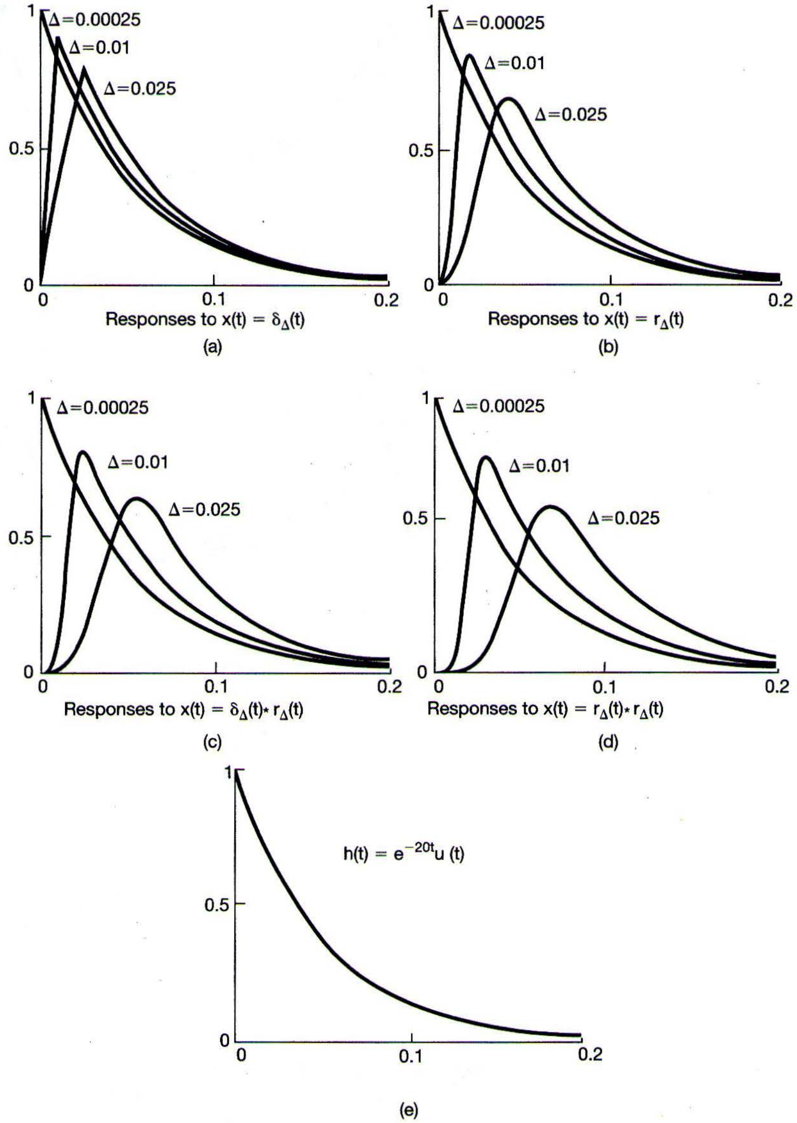


Figure 2.35 Finding a value of Δ that is "small enough" depends upon the system to which we are applying inputs: (a) responses of the causal LTI system described by eq. (2.137) to the input $\delta_\Delta(t)$ for $\Delta = 0.025$, 0.01, and 0.00025; (b) responses to $r_\Delta(t)$; (c) responses to $\delta_\Delta(t) * r_\Delta(t)$; (d) responses to $r_\Delta(t) * r_\Delta(t)$; (e) the impulse response $h(t) = e^{-20t}u(t)$ for the system. Comparing these responses to those in Figure 2.34, we see that we need to use a smaller value of Δ in this case before the duration and shape of the pulse are of no consequence.

values of Δ for the causal LTI system described by the first-order differential equation

$$\frac{dy(t)}{dt} + 20y(t) = x(t). \tag{2.137}$$

As seen in the figure, we need a smaller value of Δ in this case in order for the responses to be indistinguishable from each other and from the impulse response $h(t) = e^{-20t}u(t)$ for the system. Thus, while what we mean by "Δ small enough" is different for these two systems, we can find values of Δ small enough for both. The unit impulse is then the idealization of a short pulse whose duration is short enough for *all* systems.

2.5.2 Defining the Unit Impulse through Convolution

As the preceding example illustrates, for Δ small enough, the signals $\delta_\Delta(t)$, $r_\Delta(t)$, $r_\Delta(t) * \delta_\Delta(t)$, and $r_\Delta(t) * r_\Delta(t)$ all act like impulses when applied to an LTI system. In fact, there are many other signals for which this is true as well. What it suggests is that we should think of a unit impulse in terms of how an LTI system responds to it. While usually a function or signal is defined by what it is at each value of the independent variable, the primary importance of the unit impulse is not what it *is* at each value of t, but rather what it *does* under convolution. Thus, from the point of view of linear systems analysis, we may alternatively *define* the unit impulse as that signal which, when applied to an LTI system, yields the impulse response. That is, we define $\delta(t)$ as the signal for which

$$x(t) = x(t) * \delta(t) \tag{2.138}$$

for any $x(t)$. In this sense, signals, such as $\delta_\Delta(t)$, $r_\Delta(t)$, etc., which correspond to short pulses with vanishingly small duration as $\Delta \to 0$, all behave like a unit impulse in the limit because, if we replace $\delta(t)$ by any of these signals, then eq. (2.138) is satisfied in the limit.

All the properties of the unit impulse that we need can be obtained from the *operational definition* given by eq. (2.138). For example, if we let $x(t) = 1$ for all t, then

$$\begin{aligned} 1 = x(t) = x(t) * \delta(t) = \delta(t) * x(t) &= \int_{-\infty}^{+\infty} \delta(\tau)x(t-\tau)\,d\tau \\ &= \int_{-\infty}^{+\infty} \delta(\tau)\,d\tau, \end{aligned}$$

so that the unit impulse has unit area.

It is sometimes useful to use another completely equivalent operational definition of $\delta(t)$. To obtain this alternative form, consider taking an arbitrary signal $g(t)$, reversing it in time to obtain $g(-t)$, and then convolving this with $\delta(t)$. Using eq. (2.138), we obtain

$$g(-t) = g(-t) * \delta(t) = \int_{-\infty}^{+\infty} g(\tau - t)\,\delta(\tau)\,d\tau,$$

which, for $t = 0$, yields

$$g(0) = \int_{-\infty}^{+\infty} g(\tau)\,\delta(\tau)\,d\tau. \tag{2.139}$$

Therefore, the operational definition of $\delta(t)$ given by eq. (2.138) implies eq. (2.139). On the other hand, eq. (2.139) implies eq. (2.138). To see this, let $x(t)$ be a given signal, fix a time t, and define

$$g(\tau) = x(t-\tau).$$

Then, using eq. (2.139), we have

$$x(t) = g(0) = \int_{-\infty}^{+\infty} g(\tau)\,\delta(\tau)\,d\tau = \int_{-\infty}^{+\infty} x(t-\tau)\,\delta(\tau)\,d\tau,$$

which is precisely eq. (2.138). Therefore, eq. (2.139) is an equivalent operational definition of the unit impulse. That is, the unit impulse is the signal which, when multiplied by a signal $g(t)$ and then integrated from $-\infty$ to $+\infty$, produces the value $g(0)$.

Since we will be concerned principally with LTI systems, and thus with convolution, the characterization of $\delta(t)$ given in eq. (2.138) will be the one to which we will refer most often. However, eq. (2.139) is useful in determining some of the other properties of the unit impulse. For example, consider the signal $f(t)\,\delta(t)$, where $f(t)$ is another signal. Then, from eq. (2.139),

$$\int_{-\infty}^{+\infty} g(\tau) f(\tau)\,\delta(\tau)\,d\tau = g(0) f(0). \tag{2.140}$$

On the other hand, if we consider the signal $f(0)\,\delta(t)$, we see that

$$\int_{-\infty}^{+\infty} g(\tau) f(0)\,\delta(\tau)\,d\tau = g(0) f(0). \tag{2.141}$$

Comparing eqs. (2.140) and (2.141), we find that the two signals $f(t)\,\delta(t)$ and $f(0)\,\delta(t)$ behave identically when they are multiplied by any signal $g(t)$ and then integrated from $-\infty$ to $+\infty$. Consequently, using this form of the operational definition of signals, we conclude that

$$f(t)\,\delta(t) = f(0)\,\delta(t), \tag{2.142}$$

which is a property that we derived by alternative means in Section 1.4.2. [See eq. (1.76).]

2.5.3 Unit Doublets and Other Singularity Functions

The unit impulse is one of a class of signals known as *singularity functions,* each of which can be defined operationally in terms of its behavior under convolution. Consider the LTI system for which the output is the derivative of the input, i.e.,

$$y(t) = \frac{dx(t)}{dt} \tag{2.143}$$

The unit impulse response of this system is the derivative of the unit impulse, which is called the *unit doublet* $u_1(t)$. From the convolution representation for LTI systems, we have

$$\frac{dx(t)}{dt} = x(t) * u_1(t) \tag{2.144}$$

for any signal $x(t)$. Just as eq. (2.138) serves as the operational definition of $\delta(t)$, we will take eq. (2.144) as the operational definition of $u_1(t)$. Similarly, we can define $u_2(t)$, the second derivative of $\delta(t)$, as the impulse response of an LTI system that takes the second derivative of the input, i.e.,

$$\frac{d^2x(t)}{dt^2} = x(t) * u_2(t). \tag{2.145}$$

From eq. (2.144), we see that

$$\frac{d^2x(t)}{dt^2} = \frac{d}{dt}\left(\frac{dx(t)}{dt}\right) = x(t) * u_1(t) * u_1(t), \tag{2.146}$$

and therefore,

$$u_2(t) = u_1(t) * u_1(t). \tag{2.147}$$

In general, $u_k(t)$, $k > 0$, is the kth derivative of $\delta(t)$ and thus is the impulse response of a system that takes the kth derivative of the input. Since this system can be obtained as the cascade of k differentiators, we have

$$u_k(t) = \underbrace{u_1(t) * \cdots * u_1(t)}_{k \text{ times}}. \tag{2.148}$$

As with the unit impulse, each of these singularity functions has properties that can be derived from its operational definition. For example, if we consider the constant signal $x(t) = 1$, we find that

$$\begin{aligned} 0 = \frac{dx(t)}{dt} &= x(t) * u_1(t) = \int_{-\infty}^{+\infty} u_1(\tau)x(t-\tau)\,d\tau \\ &= \int_{-\infty}^{+\infty} u_1(\tau)\,d\tau, \end{aligned}$$

so that the unit doublet has zero area. Moreover, if we convolve the signal $g(-t)$ with $u_1(t)$, we obtain

$$\int_{-\infty}^{+\infty} g(\tau - t)u_1(\tau)\,d\tau = g(-t) * u_1(t) = \frac{dg(-t)}{dt} = -g'(-t),$$

which, for $t = 0$, yields

$$-g'(0) = \int_{-\infty}^{+\infty} g(\tau)u_1(\tau)\,d\tau. \tag{2.149}$$

In an analogous manner, we can derive related properties of $u_1(t)$ and higher order singularity functions, and several of these properties are considered in Problem 2.69.

As with the unit impulse, each of these singularity functions can be informally related to short pulses. For example, since the unit doublet is formally the derivative of the unit impulse, we can think of the doublet as the idealization of the derivative of a short pulse with unit area. For instance, consider the short pulse $\delta_\Delta(t)$ in Figure 1.34. This pulse behaves like an impulse as $\Delta \rightarrow 0$. Consequently, we would expect its derivative to behave like a doublet as $\Delta \rightarrow 0$. As verified in Problem 2.72, $d\delta_\Delta(t)/dt$ is as depicted in Figure 2.36: It consists of a unit impulse at $t = 0$ with area $+1/\Delta$, followed by a unit impulse of area $-1/\Delta$ at $t = \Delta$, i.e.,

$$\frac{d\delta_\Delta(t)}{dt} = \frac{1}{\Delta}[\delta(t) - \delta(t - \Delta)]. \tag{2.150}$$

Consequently, using the fact that $x(t) * \delta(t - t_0) = x(t - t_0)$ [see eq. (2.70)], we find that

$$x(t) * \frac{d\delta_\Delta(t)}{dt} = \frac{x(t) - x(t - \Delta)}{\Delta} \cong \frac{dx(t)}{dt}, \tag{2.151}$$

where the approximation becomes increasingly accurate as $\Delta \rightarrow 0$. Comparing eq. (2.151) with eq. (2.144), we see that $d\delta_\Delta(t)/dt$ does indeed behave like a unit doublet as $\Delta \rightarrow 0$.

In addition to singularity functions that are derivatives of different orders of the unit impulse, we can also define signals that represent successive integrals of the unit impulse function. As we saw in Example 2.13, the unit step is the impulse response of an integrator:

$$y(t) = \int_{-\infty}^{t} x(\tau)\,d\tau.$$

Therefore,

$$u(t) = \int_{-\infty}^{t} \delta(\tau)\,d\tau, \tag{2.152}$$

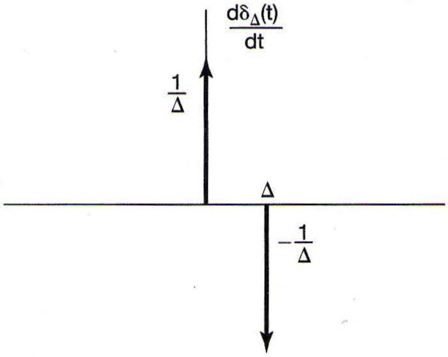


Figure 2.36 The derivative $d\delta_\Delta(t)/dt$ of the short rectangular pulse $\delta_\Delta(t)$ of Figure 1.34.

and we also have the following operational definition of $u(t)$:

$$x(t) * u(t) = \int_{-\infty}^{t} x(\tau)\, d\tau. \tag{2.153}$$

Similarly, we can define the system that consists of a cascade of two integrators. Its impulse response is denoted by $u_{-2}(t)$, which is simply the convolution of $u(t)$, the impulse response of one integrator, with itself:

$$u_{-2}(t) = u(t) * u(t) = \int_{-\infty}^{t} u(\tau)\, d\tau. \tag{2.154}$$

Since $u(t)$ equals 0 for $t < 0$ and equals 1 for $t > 0$, it follows that

$$u_{-2}(t) = tu(t). \tag{2.155}$$

This signal, which is referred to as the *unit ramp function*, is shown in Figure 2.37. Also, we can obtain an operational definition for the behavior of $u_{-2}(t)$ under convolution from eqs. (2.153) and (2.154):

$$\begin{aligned} x(t) * u_{-2}(t) &= x(t) * u(t) * u(t) \\ &= \left(\int_{-\infty}^{t} x(\sigma)\, d\sigma\right) * u(t) \\ &= \int_{-\infty}^{t} \left(\int_{-\infty}^{\tau} x(\sigma)\, d\sigma\right) d\tau. \end{aligned} \tag{2.156}$$

In an analogous fashion, we can define higher order integrals of $\delta(t)$ as the impulse responses of cascades of integrators:

$$u_{-k}(t) = \underbrace{u(t) * \cdots * u(t)}_{k \text{ times}} = \int_{-\infty}^{t} u_{-(k-1)}(\tau)\, d\tau. \tag{2.157}$$

The convolution of $x(t)$ with $u_{-3}(t), u_{-4}(t), \ldots$ generate correspondingly higher order integrals of $x(t)$. Also, note that the integrals in eq. (2.157) can be evaluated directly (see

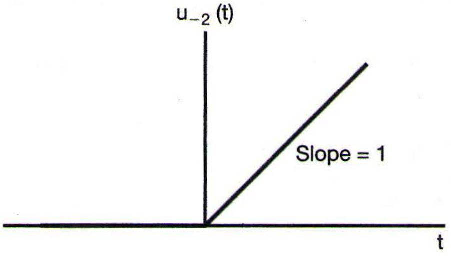


Figure 2.37 Unit ramp function.

Problem 2.73), as was done in eq. (2.155), to obtain

$$u_{-k}(t) = \frac{t^{k-1}}{(k-1)!}u(t). \tag{2.158}$$

Thus, unlike the derivatives of $\delta(t)$, the successive integrals of the unit impulse are functions that can be defined for each value of t [eq. (2.158)], as well as by their behavior under convolution.

At times it will be worthwhile to use an alternative notation for $\delta(t)$ and $u(t)$, namely,

$$\delta(t) = u_0(t), \tag{2.159}$$

$$u(t) = u_{-1}(t). \tag{2.160}$$

With this notation, $u_k(t)$ for $k > 0$ denotes the impulse response of a cascade of k differentiators, $u_0(t)$ is the impulse response of the identity system, and, for $k < 0$, $u_k(t)$ is the impulse response of a cascade of $|k|$ integrators. Furthermore, since a differentiator is the inverse system of an integrator,

$$u(t) * u_1(t) = \delta(t),$$

or, in our alternative notation,

$$u_{-1}(t) * u_1(t) = u_0(t). \tag{2.161}$$

More generally, from eqs. (2.148), (2.157), and (2.161), we see that for any integers k and r,

$$u_k(t) * u_r(t) = u_{k+r}(t). \tag{2.162}$$

If k and r are both positive, eq. (2.162) states that a cascade of k differentiators followed by r more differentiators yields an output that is the $(k+r)$th derivative of the input. Similarly, if k is negative and r is negative, we have a cascade of $|k|$ integrators followed by another $|r|$ integrators. Also, if k is negative and r is positive, we have a cascade of $|k|$ integrators followed by r differentiators, and the overall system is equivalent to a cascade of $|k + r|$ integrators if $k+r < 0$, a cascade of $k+r$ differentiators if $k+r > 0$, or the identity system if $k+r = 0$. Therefore, by defining singularity functions in terms of their behavior under convolution, we obtain a characterization that allows us to manipulate them with relative ease and to interpret them directly in terms of their significance for LTI systems. Since this is our primary concern in the book, the operational definition for singularity functions that we have given in this section will suffice for our purposes.[6]

---

[6] As mentioned in Chapter 1, singularity functions have been heavily studied in the field of mathematics under the alternative names of *generalized functions* and *distribution theory*. The approach we have taken in this section is actually closely allied in spirit with the rigorous approach taken in the references given in footnote 3 of Section 1.4.

2.6 SUMMARY

In this chapter, we have developed important representations for LTI systems, both in discrete time and in continuous time. In discrete time we derived a representation of signals as weighted sums of shifted unit impulses, and we then used this to derive the convolution-sum representation for the response of a discrete-time LTI system. In continuous time we derived an analogous representation of continuous-time signals as weighted integrals of shifted unit impulses, and we used this to derive the convolution integral representation for continuous-time LTI systems. These representations are extremely important, as they allow us to compute the response of an LTI system to an arbitrary input in terms of the system's response to a unit impulse. Moreover, in Section 2.3 the convolution sum and integral provided us with a means of analyzing the properties of LTI systems and, in particular, of relating LTI system properties, including causality and stability, to corresponding properties of the unit impulse response. Also, in Section 2.5 we developed an interpretation of the continuous-time unit impulse and other related singularity functions in terms of their behavior under convolution. This interpretation is particularly useful in the analysis of LTI systems.

An important class of continuous-time systems consists of those described by linear constant-coefficient differential equations. Similarly, in discrete time, linear constant-coefficient difference equations play an equally important role. In Section 2.4, we examined simple examples of differential and difference equations and discussed some of the properties of systems described by these types of equations. In particular, systems described by linear constant-coefficient differential and difference equations together with the condition of initial rest are causal and LTI. In subsequent chapters, we will develop additional tools that greatly facilitate our ability to analyze such systems.

Chapter 2 Problems

The first section of problems belongs to the basic category, and the answers are provided in the back of the book. The remaining three sections contain problems belonging to the basic, advanced, and extension categories, respectively.

Extension problems introduce applications, concepts, or methods beyond those presented in the text.

BASIC PROBLEMS WITH ANSWERS

2.1. Let

$$x[n] = \delta[n] + 2\delta[n-1] - \delta[n-3] \quad \text{and} \quad h[n] = 2\delta[n+1] + 2\delta[n-1].$$

Compute and plot each of the following convolutions:

(a) $y_1[n] = x[n] * h[n]$ **(b)** $y_2[n] = x[n+2] * h[n]$

(c) $y_3[n] = x[n] * h[n+2]$

2.2. Consider the signal

$$h[n] = \left(\frac{1}{2}\right)^{n-1} \{u[n+3] - u[n-10]\}.$$

Express A and B in terms of n so that the following equation holds:

$$h[n-k] = \begin{cases} (\frac{1}{2})^{n-k-1}, & A \leq k \leq B \\ 0, & \text{elsewhere} \end{cases}.$$

2.3. Consider an input $x[n]$ and a unit impulse response $h[n]$ given by

$$x[n] = \left(\frac{1}{2}\right)^{n-2} u[n-2],$$

$$h[n] = u[n+2].$$

Determine and plot the output $y[n] = x[n] * h[n]$.

2.4. Compute and plot $y[n] = x[n] * h[n]$, where

$$x[n] = \begin{cases} 1, & 3 \leq n \leq 8 \\ 0, & \text{otherwise} \end{cases},$$

$$h[n] = \begin{cases} 1, & 4 \leq n \leq 15 \\ 0, & \text{otherwise} \end{cases}.$$

2.5. Let

$$x[n] = \begin{cases} 1, & 0 \leq n \leq 9 \\ 0, & \text{elsewhere} \end{cases} \quad \text{and} \quad h[n] = \begin{cases} 1, & 0 \leq n \leq N \\ 0, & \text{elsewhere} \end{cases},$$

where $N \leq 9$ is an integer. Determine the value of N, given that $y[n] = x[n] * h[n]$ and

$$y[4] = 5, \quad y[14] = 0.$$

2.6. Compute and plot the convolution $y[n] = x[n] * h[n]$, where

$$x[n] = \left(\frac{1}{3}\right)^{-n} u[-n-1] \quad \text{and} \quad h[n] = u[n-1].$$

2.7. A linear system S has the relationship

$$y[n] = \sum_{k=-\infty}^{\infty} x[k]g[n-2k]$$

between its input $x[n]$ and its output $y[n]$, where $g[n] = u[n] - u[n-4]$.

(a) Determine $y[n]$ when $x[n] = \delta[n-1]$.
(b) Determine $y[n]$ when $x[n] = \delta[n-2]$.
(c) Is S LTI?
(d) Determine $y[n]$ when $x[n] = u[n]$.

2.8. Determine and sketch the convolution of the following two signals:

$$x(t) = \begin{cases} t+1, & 0 \le t \le 1 \\ 2-t, & 1 < t \le 2 \\ 0, & \text{elsewhere} \end{cases},$$

$$h(t) = \delta(t+2) + 2\delta(t+1).$$

2.9. Let

$$h(t) = e^{2t}u(-t+4) + e^{-2t}u(t-5).$$

Determine A and B such that

$$h(t-\tau) = \begin{cases} e^{-2(t-\tau)}, & \tau < A \\ 0, & A < \tau < B \\ e^{2(t-\tau)}, & B < \tau \end{cases}.$$

2.10. Suppose that

$$x(t) = \begin{cases} 1, & 0 \le t \le 1 \\ 0, & \text{elsewhere} \end{cases}$$

and $h(t) = x(t/\alpha)$, where $0 < \alpha \le 1$.
(a) Determine and sketch $y(t) = x(t) * h(t)$.
(b) If $dy(t)/dt$ contains only three discontinuities, what is the value of α?

2.11. Let

$$x(t) = u(t-3) - u(t-5) \quad \text{and} \quad h(t) = e^{-3t}u(t).$$

(a) Compute $y(t) = x(t) * h(t)$.
(b) Compute $g(t) = (dx(t)/dt) * h(t)$.
(c) How is $g(t)$ related to $y(t)$?

2.12. Let

$$y(t) = e^{-t}u(t) * \sum_{k=-\infty}^{\infty} \delta(t-3k).$$

Show that $y(t) = Ae^{-t}$ for $0 \le t < 3$, and determine the value of A.

2.13. Consider a discrete-time system S_1 with impulse response

$$h[n] = \left(\frac{1}{5}\right)^n u[n].$$

(a) Find the integer A such that $h[n] - Ah[n-1] = \delta[n]$.
(b) Using the result from part (a), determine the impulse response $g[n]$ of an LTI system S_2 which is the inverse system of S_1.

2.14. Which of the following impulse responses correspond(s) to stable LTI systems?
(a) $h_1(t) = e^{-(1-2j)t}u(t)$ **(b)** $h_2(t) = e^{-t}\cos(2t)u(t)$

2.15. Which of the following impulse responses correspond(s) to stable LTI systems?
(a) $h_1[n] = n\cos(\frac{\pi}{4}n)u[n]$ **(b)** $h_2[n] = 3^n u[-n+10]$

2.16. For each of the following statements, determine whether it is true or false:
(a) If $x[n] = 0$ for $n < N_1$ and $h[n] = 0$ for $n < N_2$, then $x[n] * h[n] = 0$ for $n < N_1 + N_2$.
(b) If $y[n] = x[n] * h[n]$, then $y[n-1] = x[n-1] * h[n-1]$.
(c) If $y(t) = x(t) * h(t)$, then $y(-t) = x(-t) * h(-t)$.
(d) If $x(t) = 0$ for $t > T_1$ and $h(t) = 0$ for $t > T_2$, then $x(t) * h(t) = 0$ for $t > T_1 + T_2$.

2.17. Consider an LTI system whose input $x(t)$ and output $y(t)$ are related by the differential equation

$$\frac{d}{dt}y(t) + 4y(t) = x(t). \qquad \text{(P2.17–1)}$$

The system also satisfies the condition of initial rest.
(a) If $x(t) = e^{(-1+3j)t}u(t)$, what is $y(t)$?
(b) Note that $\mathcal{R}e\{x(t)\}$ will satisfy eq. (P2.17–1) with $\mathcal{R}e\{y(t)\}$. Determine the output $y(t)$ of the LTI system if

$$x(t) = e^{-t}\cos(3t)u(t).$$

2.18. Consider a causal LTI system whose input $x[n]$ and output $y[n]$ are related by the difference equation

$$y[n] = \frac{1}{4}y[n-1] + x[n].$$

Determine $y[n]$ if $x[n] = \delta[n-1]$.

2.19. Consider the cascade of the following two systems S_1 and S_2, as depicted in Figure P2.19:

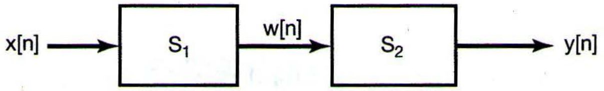


Figure P2.19

$$\begin{aligned} &S_1 : \text{causal LTI,} \\ &\qquad w[n] = \frac{1}{2}w[n-1] + x[n]; \\ &S_2 : \text{causal LTI,} \\ &\qquad y[n] = \alpha y[n-1] + \beta w[n]. \end{aligned}$$

The difference equation relating $x[n]$ and $y[n]$ is:

$$y[n] = -\frac{1}{8}y[n-2] + \frac{3}{4}y[n-1] + x[n].$$

(a) Determine α and β.
(b) Show the impulse response of the cascade connection of S_1 and S_2.

2.20. Evaluate the following integrals:

(a) $\int_{-\infty}^{\infty} u_0(t)\cos(t)\,dt$
(b) $\int_0^5 \sin(2\pi t)\,\delta(t+3)\,dt$
(c) $\int_{-5}^{5} u_1(1-\tau)\cos(2\pi\tau)\,d\tau$

BASIC PROBLEMS

2.21. Compute the convolution $y[n] = x[n] * h[n]$ of the following pairs of signals:

(a) $\left.\begin{aligned} x[n] &= \alpha^n u[n], \\ h[n] &= \beta^n u[n], \end{aligned}\right\} \quad \alpha \neq \beta$
(b) $x[n] = h[n] = \alpha^n u[n]$
(c) $x[n] = (-\frac{1}{2})^n u[n-4]$
$h[n] = 4^n u[2 - n]$
(d) $x[n]$ and $h[n]$ are as in Figure P2.21.

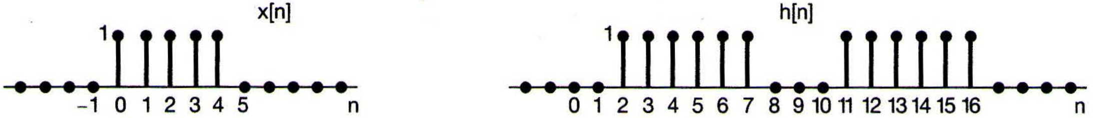


Figure P2.21

2.22. For each of the following pairs of waveforms, use the convolution integral to find the response $y(t)$ of the LTI system with impulse response $h(t)$ to the input $x(t)$. Sketch your results.

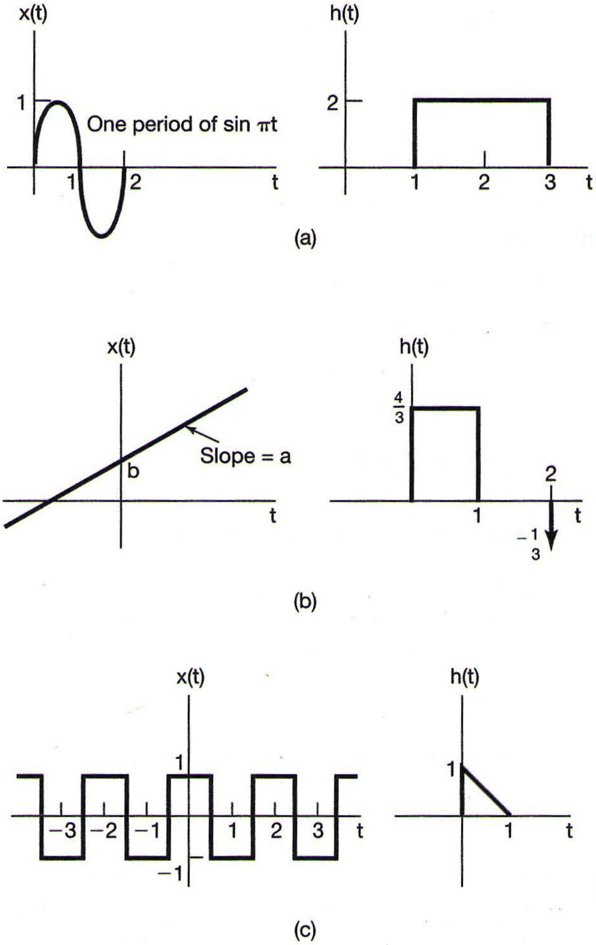
(a) $\left.\begin{aligned} x(t) &= e^{-\alpha t}u(t) \\ h(t) &= e^{-\beta t}u(t) \end{aligned}\right\}$ (Do this both when $\alpha \neq \beta$ and when $\alpha = \beta$.)

(b) $x(t) = u(t) - 2u(t-2) + u(t-5)$
$h(t) = e^{2t}u(1-t)$

(c) $x(t)$ and $h(t)$ are as in Figure P2.22(a).

(d) $x(t)$ and $h(t)$ are as in Figure P2.22(b).

(e) $x(t)$ and $h(t)$ are as in Figure P2.22(c).


x(t)
1
One period of sin πt
1
2
t
h(t)
2
1
2
3
t
x(t)
b
Slope = a
t
h(t)
4/3
2
1
−1/3
t
x(t)
1
−3
−2
−1
1
2
3
t
−1
h(t)
1
1
t

(a)

(b)

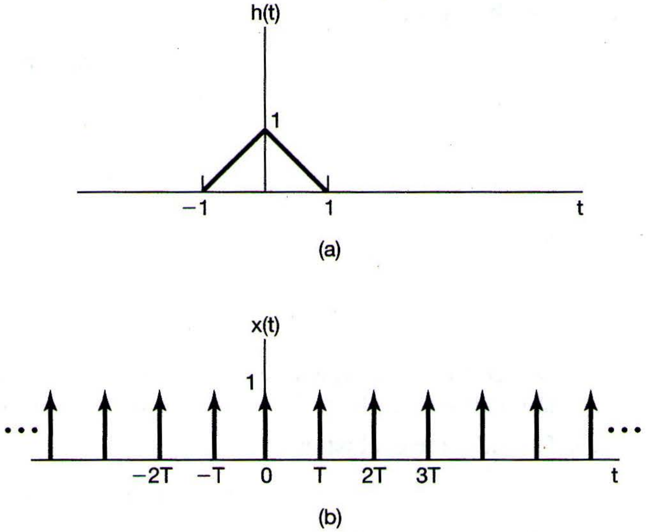
(c)

Figure P2.22

2.23. Let $h(t)$ be the triangular pulse shown in Figure P2.23(a), and let $x(t)$ be the impulse train depicted in Figure P2.23(b). That is,

$$x(t) = \sum_{k=-\infty}^{+\infty} \delta(t - kT).$$

Determine and sketch $y(t) = x(t) * h(t)$ for the following values of T:
(a) $T = 4$ **(b)** $T = 2$ **(c)** $T = 3/2$ **(d)** $T = 1$


h(t)
1
−1
1
t
x(t)
1
• • •
• • •
−2T
−T
0
T
2T
3T
t

(a)

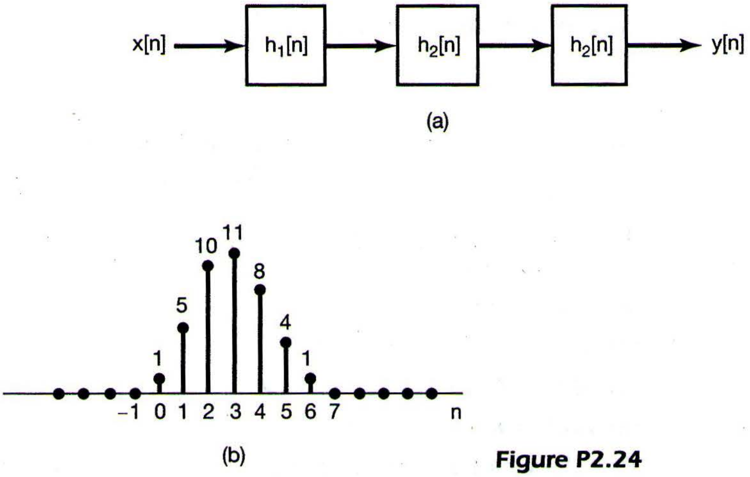
(b)

Figure P2.23

2.24. Consider the cascade interconnection of three causal LTI systems, illustrated in Figure P2.24(a). The impulse response $h_2[n]$ is

$$h_2[n] = u[n] - u[n-2],$$

and the overall impulse response is as shown in Figure P2.24(b).


x[n]
h1[n]
h2[n]
h2[n]
y[n]
10
11
8
5
4
1
1
-1
0
1
2
3
4
5
6
7
n

(a)

(b)

Figure P2.24

(a) Find the impulse response $h_1[n]$.
(b) Find the response of the overall system to the input

$$x[n] = \delta[n] - \delta[n-1].$$

2.25. Let the signal

$$y[n] = x[n] * h[n],$$

where

$$x[n] = 3^n u[-n-1] + \left(\frac{1}{3}\right)^n u[n]$$

and

$$h[n] = \left(\frac{1}{4}\right)^n u[n+3].$$

(a) Determine $y[n]$ *without* utilizing the distributive property of convolution.
(b) Determine $y[n]$ *utilizing* the distributive property of convolution.

2.26. Consider the evaluation of

$$y[n] = x_1[n] * x_2[n] * x_3[n],$$

where $x_1[n] = (0.5)^n u[n]$, $x_2[n] = u[n+3]$, and $x_3[n] = \delta[n] - \delta[n-1]$.
(a) Evaluate the convolution $x_1[n] * x_2[n]$.
(b) Convolve the result of part (a) with $x_3[n]$ in order to evaluate $y[n]$.
(c) Evaluate the convolution $x_2[n] * x_3[n]$.
(d) Convolve the result of part (c) with $x_1[n]$ in order to evaluate $y[n]$.

2.27. We define the area under a continuous-time signal $v(t)$ as

$$A_v = \int_{-\infty}^{+\infty} v(t)\,dt.$$

Show that if $y(t) = x(t) * h(t)$, then

$$A_y = A_x A_h.$$

2.28. The following are the impulse responses of discrete-time LTI systems. Determine whether each system is causal and/or stable. Justify your answers.
(a) $h[n] = (\frac{1}{5})^n u[n]$
(b) $h[n] = (0.8)^n u[n+2]$
(c) $h[n] = (\frac{1}{2})^n u[-n]$
(d) $h[n] = (5)^n u[3-n]$
(e) $h[n] = (-\frac{1}{2})^n u[n] + (1.01)^n u[n-1]$
(f) $h[n] = (-\frac{1}{2})^n u[n] + (1.01)^n u[1-n]$
(g) $h[n] = n(\frac{1}{3})^n u[n-1]$

2.29. The following are the impulse responses of continuous-time LTI systems. Determine whether each system is causal and/or stable. Justify your answers.
(a) $h(t) = e^{-4t} u(t-2)$
(b) $h(t) = e^{-6t} u(3-t)$
(c) $h(t) = e^{-2t} u(t+50)$
(d) $h(t) = e^{2t} u(-1-t)$

(e) $h(t) = e^{-6|t|}$
(f) $h(t) = te^{-t}u(t)$
(g) $h(t) = (2e^{-t} - e^{(t-100)/100})u(t)$

2.30. Consider the first-order difference equation

$$y[n] + 2y[n-1] = x[n].$$

Assuming the condition of initial rest (i.e., if $x[n] = 0$ for $n < n_0$, then $y[n] = 0$ for $n < n_0$), find the impulse response of a system whose input and output are related by this difference equation. You may solve the problem by rearranging the difference equation so as to express $y[n]$ in terms of $y[n-1]$ and $x[n]$ and generating the values of $y[0],\ y[+1],\ y[+2], \ldots$ in that order.

2.31. Consider the LTI system initially at rest and described by the difference equation

$$y[n] + 2y[n-1] = x[n] + 2x[n-2].$$

Find the response of this system to the input depicted in Figure P2.31 by solving the difference equation recursively.

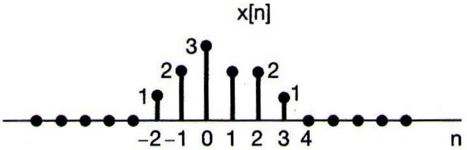


Figure P2.31

2.32. Consider the difference equation

$$y[n] - \frac{1}{2}y[n-1] = x[n], \tag{P2.32–1}$$

and suppose that

$$x[n] = \left(\frac{1}{3}\right)^n u[n]. \tag{P2.32–2}$$

Assume that the solution $y[n]$ consists of the sum of a particular solution $y_p[n]$ to eq. (P2.32–1) and a homogeneous solution $y_h[n]$ satisfying the equation

$$y_h[n] - \frac{1}{2}y_h[n-1] = 0.$$

(a) Verify that the homogeneous solution is given by

$$y_h[n] = A\left(\frac{1}{2}\right)^n$$

(b) Let us consider obtaining a particular solution $y_p[n]$ such that

$$y_p[n] - \frac{1}{2}y_p[n-1] = \left(\frac{1}{3}\right)^n u[n].$$

By assuming that $y_p[n]$ is of the form $B(\frac{1}{3})^n$ for $n \geq 0$, and substituting this in the above difference equation, determine the value of B.

(c) Suppose that the LTI system described by eq. (P2.32–1) and initially at rest has as its input the signal specified by eq. (P2.32–2). Since $x[n] = 0$ for $n < 0$, we have that $y[n] = 0$ for $n < 0$. Also, from parts (a) and (b) we have that $y[n]$ has the form

$$y[n] = A\left(\frac{1}{2}\right)^n + B\left(\frac{1}{3}\right)^n$$

for $n \geq 0$. In order to solve for the unknown constant A, we must specify a value for $y[n]$ for some $n \geq 0$. Use the condition of initial rest and eqs. (P2.32–1) and (P2.32–2) to determine $y[0]$. From this value determine the constant A. The result of this calculation yields the solution to the difference equation (P2.32–1) under the condition of initial rest, when the input is given by eq. (P2.32–2).

2.33. Consider a system whose input $x(t)$ and output $y(t)$ satisfy the first-order differential equation

$$\frac{dy(t)}{dt} + 2y(t) = x(t). \tag{P2.33–1}$$

The system also satisfies the condition of initial rest.

(a) (i) Determine the system output $y_1(t)$ when the input is $x_1(t) = e^{3t}u(t)$.

(ii) Determine the system output $y_2(t)$ when the input is $x_2(t) = e^{2t}u(t)$.

(iii) Determine the system output $y_3(t)$ when the input is $x_3(t) = \alpha e^{3t}u(t) + \beta e^{2t}u(t)$, where α and β are real numbers. Show that $y_3(t) = \alpha y_1(t) + \beta y_2(t)$.

(iv) Now let $x_1(t)$ and $x_2(t)$ be arbitrary signals such that

$$x_1(t) = 0, \ \text{for } t < t_1,$$
$$x_2(t) = 0, \ \text{for } t < t_2.$$

Letting $y_1(t)$ be the system output for input $x_1(t)$, $y_2(t)$ be the system output for input $x_2(t)$, and $y_3(t)$ be the system output for $x_3(t) = \alpha x_1(t) + \beta x_2(t)$, show that

$$y_3(t) = \alpha y_1(t) + \beta y_2(t).$$

We may therefore conclude that the system under consideration is linear.

(b) (i) Determine the system output $y_1(t)$ when the input is $x_1(t) = Ke^{2t}u(t)$.

(ii) Determine the system output $y_2(t)$ when the input is $x_2(t) = Ke^{2(t-T)}u(t-T)$. Show that $y_2(t) = y_1(t-T)$.

(iii) Now let $x_1(t)$ be an arbitrary signal such that $x_1(t) = 0$ for $t < t_0$. Letting $y_1(t)$ be the system output for input $x_1(t)$ and $y_2(t)$ be the system output for $x_2(t) = x_1(t-T)$, show that

$$y_2(t) = y_1(t-T).$$

We may therefore conclude that the system under consideration is time invariant. In conjunction with the result derived in part (a), we conclude that the given system is LTI. Since this system satisfies the condition of initial rest, it is causal as well.

2.34. The initial rest assumption corresponds to a zero-valued auxiliary condition being imposed at a time determined in accordance with the input signal. In this problem we show that if the auxiliary condition used is nonzero or if it is always applied at a fixed time (regardless of the input signal) the corresponding system cannot be LTI. Consider a system whose input $x(t)$ and output $y(t)$ satisfy the first-order differential equation (P2.33–1).

(a) Given the auxiliary condition $y(1) = 1$, use a counterexample to show that the system is not linear.

(b) Given the auxiliary condition $y(1) = 1$, use a counterexample to show that the system is not time invariant.

(c) Given the auxiliary condition $y(1) = 1$, show that the system is incrementally linear.

(d) Given the auxiliary condition $y(1) = 0$, show that the system is linear but not time invariant.

(e) Given the auxiliary condition $y(0) + y(4) = 0$, show that the system is linear but not time invariant.

2.35. In the previous problem we saw that application of an auxiliary condition at a fixed time (regardless of the input signal) leads to the corresponding system being not time-invariant. In this problem, we explore the effect of fixed auxiliary conditions on the causality of a system. Consider a system whose input $x(t)$ and output $y(t)$ satisfy the first-order differential equation (P2.33–1). Assume that the auxiliary condition associated with the differential equation is $y(0) = 0$. Determine the output of the system for each of the following two inputs:

(a) $x_1(t) = 0, \text{ for all } t$

(b) $x_2(t) = \begin{cases} 0, & t < -1 \\ 1, & t > -1 \end{cases}$

Observe that if $y_1(t)$ is the output for input $x_1(t)$ and $y_2(t)$ is the output for input $x_2(t)$, then $y_1(t)$ and $y_2(t)$ are not identical for $t < -1$, even though $x_1(t)$ and $x_2(t)$ are identical for $t < -1$. Use this observation as the basis of an argument to conclude that the given system is not causal.

2.36. Consider a discrete-time system whose input $x[n]$ and output $y[n]$ are related by

$$y[n] = \left(\frac{1}{2}\right)y[n-1] + x[n].$$

(a) Show that if this system satisfies the condition of initial rest (i.e., if $x[n] = 0$ for $n < n_0$, then $y[n] = 0$ for $n < n_0$), then it is linear and time invariant.

(b) Show that if this system does not satisfy the condition of initial rest, but instead uses the auxiliary condition $y[0] = 0$, it is not causal. [*Hint:* Use an approach similar to that used in Problem 2.35.]

2.37. Consider a system whose input and output are related by the first-order differential equation (P2.33–1). Assume that the system satisfies the condition of final rest [i. e., if $x(t) = 0$ for $t > t_0$, then $y(t) = 0$ for $t > t_0$]. Show that this system is *not* causal. [*Hint:* Consider two inputs to the system, $x_1(t) = 0$ and $x_2(t) = e^t(u(t) - u(t-1))$, which result in outputs $y_1(t)$ and $y_2(t)$, respectively. Then show that $y_1(t) \neq y_2(t)$ for $t < 0$.]

2.38. Draw block diagram representations for causal LTI systems described by the following difference equations:

(a) $y[n] = \frac{1}{3}y[n-1] + \frac{1}{2}x[n]$

(b) $y[n] = \frac{1}{3}y[n-1] + x[n-1]$

2.39. Draw block diagram representations for causal LTI systems described by the following differential equations:

(a) $y(t) = -(\frac{1}{2})\,dy(t)/dt + 4x(t)$

(b) $dy(t)/dt + 3y(t) = x(t)$

ADVANCED PROBLEMS

2.40. **(a)** Consider an LTI system with input and output related through the equation

$$y(t) = \int_{-\infty}^{t} e^{-(t-\tau)}x(\tau - 2)\,d\tau.$$

What is the impulse response $h(t)$ for this system?

(b) Determine the response of the system when the input $x(t)$ is as shown in Figure P2.40.

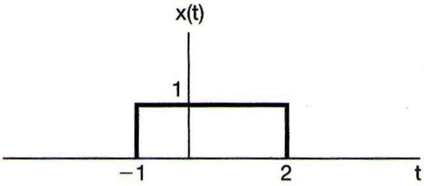


Figure P2.40

2.41. Consider the signal

$$x[n] = \alpha^n u[n].$$

(a) Sketch the signal $g[n] = x[n] - \alpha x[n-1]$.

(b) Use the result of part (a) in conjunction with properties of convolution in order to determine a sequence $h[n]$ such that

$$x[n] * h[n] = \left(\frac{1}{2}\right)^n \{u[n+2] - u[n-2]\}.$$

2.42. Suppose that the signal

$$x(t) = u(t + 0.5) - u(t - 0.5)$$

is convolved with the signal

$$h(t) = e^{j\omega_0 t}.$$

(a) Determine a value of ω_0 which ensures that

$$y(0) = 0,$$

where $y(t) = x(t) * h(t)$.

(b) Is your answer to the previous part unique?

2.43. One of the important properties of convolution, in both continuous and discrete time, is the associativity property. In this problem, we will check and illustrate this property.

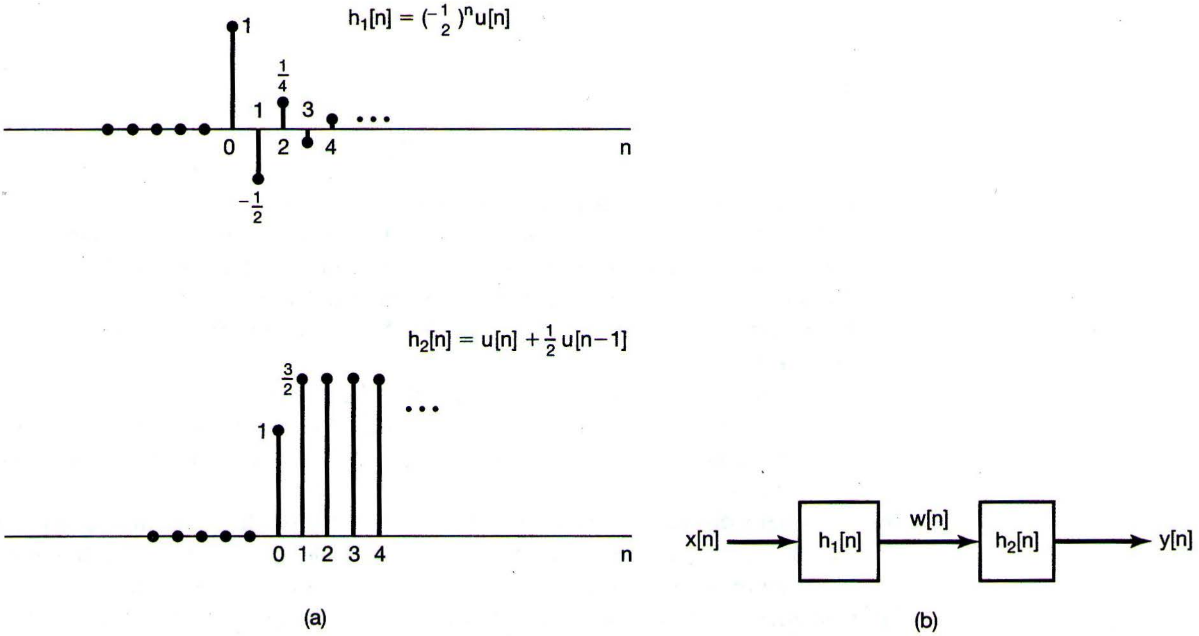
(a) Prove the equality

$$[x(t) * h(t)] * g(t) = x(t) * [h(t) * g(t)] \qquad \text{(P2.43–1)}$$

by showing that both sides of eq. (P2.43–1) equal

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x(\tau)h(\sigma)g(t-\tau-\sigma)\,d\tau\,d\sigma.$$

(b) Consider two LTI systems with the unit sample responses $h_1[n]$ and $h_2[n]$ shown in Figure P2.43(a). These two systems are cascaded as shown in Figure P2.43(b). Let $x[n] = u[n]$.



(a)

(b)

Figure P2.43

(i) Compute $y[n]$ by first computing $w[n] = x[n] * h_1[n]$ and then computing $y[n] = w[n] * h_2[n]$; that is, $y[n] = [x[n] * h_1[n]] * h_2[n]$.

(ii) Now find $y[n]$ by first convolving $h_1[n]$ and $h_2[n]$ to obtain $g[n] = h_1[n] * h_2[n]$ and then convolving $x[n]$ with $g[n]$ to obtain $y[n] = x[n] * [h_1[n] * h_2[n]]$.

The answers to (i) and (ii) should be identical, illustrating the associativity property of discrete-time convolution.

(c) Consider the cascade of two LTI systems as in Figure P2.43(b), where in this case

$$h_1[n] = \sin 8n$$

and

$$h_2[n] = a^n u[n], \quad |a| < 1,$$

and where the input is

$$x[n] = \delta[n] - a\delta[n-1].$$

Determine the output $y[n]$. (*Hint:* The use of the associative and commutative properties of convolution should greatly facilitate the solution.)

2.44. (a) If

$$x(t) = 0, \ |t| > T_1,$$

and

$$h(t) = 0, \ |t| > T_2,$$

then

$$x(t) * h(t) = 0, \ |t| > T_3$$

for some positive number T_3. Express T_3 in terms of T_1 and T_2.

(b) A discrete-time LTI system has input $x[n]$, impulse response $h[n]$, and output $y[n]$. If $h[n]$ is known to be zero everywhere outside the interval $N_0 \le n \le N_1$ and $x[n]$ is known to be zero everywhere outside the interval $N_2 \le n \le N_3$, then the output $y[n]$ is constrained to be zero everywhere, except on some interval $N_4 \le n \le N_5$.

(i) Determine N_4 and N_5 in terms of N_0, N_1, N_2, and N_3.

(ii) If the interval $N_0 \le n \le N_1$ is of length M_h, $N_2 \le n \le N_3$ is of length M_x, and $N_4 \le n \le N_5$ is of length M_y, express M_y in terms of M_h and M_x.

(c) Consider a discrete-time LTI system with the property that if the input $x[n] = 0$ for all $n \ge 10$, then the output $y[n] = 0$ for all $n \ge 15$. What condition must $h[n]$, the impulse response of the system, satisfy for this to be true?

(d) Consider an LTI system with impulse response in Figure P2.44. Over what interval must we know $x(t)$ in order to determine $y(0)$?

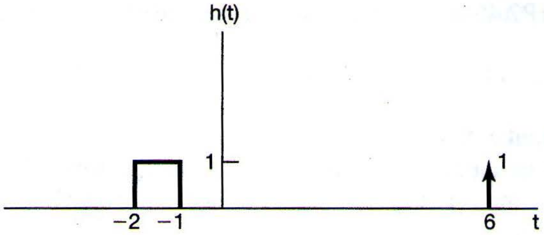


Figure P2.44

2.45. **(a)** Show that if the response of an LTI system to $x(t)$ is the output $y(t)$, then the response of the system to

$$x'(t) = \frac{dx(t)}{dt}$$

is $y'(t)$. Do this problem in three different ways:

(i) Directly from the properties of linearity and time invariance and the fact that

$$x'(t) = \lim_{h \to 0} \frac{x(t) - x(t-h)}{h}.$$

(ii) By differentiating the convolution integral.

(iii) By examining the system in Figure P2.45.

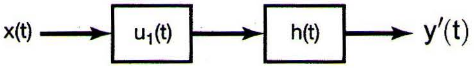


Figure P2.45

(b) Demonstrate the validity of the following relationships:

(i) $y'(t) = x(t) * h'(t)$

(ii) $y(t) = \left(\int_{-\infty}^{t} x(\tau)\,d\tau\right) * h'(t) = \int_{-\infty}^{t} [x'(\tau) * h(\tau)]d\tau = x'(t) * \left(\int_{-\infty}^{t} h(\tau)d\tau\right)$

[*Hint:* These are easily done using block diagrams as in (iii) of part (a) and the fact that $u_1(t) * u_{-1}(t) = \delta(t)$.]

(c) An LTI system has the response $y(t) = \sin \omega_0 t$ to input $x(t) = e^{-5t}u(t)$. Use the result of part (a) to aid in determining the impulse response of this system.

(d) Let $s(t)$ be the unit step response of a continuous-time LTI system. Use part (b) to deduce that the response $y(t)$ to the input $x(t)$ is

$$y(t) = \int_{-\infty}^{+\infty} x'(\tau)s(t-\tau)d\tau. \qquad \text{(P2.45–1)}$$

Show also that

$$x(t) = \int_{-\infty}^{+\infty} x'(\tau)u(t-\tau)\,d\tau. \qquad \text{(P2.45–2)}$$

(e) Use eq. (P2.45–1) to determine the response of an LTI system with step response

$$s(t) = (e^{-3t} - 2e^{-2t} + 1)u(t)$$

to the input $x(t) = e^t u(t)$.

(f) Let $s[n]$ be the unit step response of a discrete-time LTI system. What are the discrete-time counterparts of eqs. (P2.45–1) and (P2.45–2)?

2.46. Consider an LTI system S and a signal $x(t) = 2e^{-3t}u(t-1)$. If

$$x(t) \longrightarrow y(t)$$

and

$$\frac{dx(t)}{dt} \longrightarrow -3y(t) + e^{-2t}u(t),$$

determine the impulse response $h(t)$ of S.

2.47. We are given a certain linear time-invariant system with impulse response $h_0(t)$. We are told that when the input is $x_0(t)$ the output is $y_0(t)$, which is sketched in Figure P2.47. We are then given the following set of inputs to linear time-invariant systems with the indicated impulse responses:

	Input x(t)	*Impulse response h(t)*
(a)	$x(t) = 2x_0(t)$	$h(t) = h_0(t)$
(b)	$x(t) = x_0(t) - x_0(t-2)$	$h(t) = h_0(t)$
(c)	$x(t) = x_0(t-2)$	$h(t) = h_0(t+1)$
(d)	$x(t) = x_0(-t)$	$h(t) = h_0(t)$
(e)	$x(t) = x_0(-t)$	$h(t) = h_0(-t)$
(f)	$x(t) = x_0'(t)$	$h(t) = h_0'(t)$

[Here $x_0'(t)$ and $h_0'(t)$ denote the first derivatives of $x_0(t)$ and $h_0(t)$, respectively.]

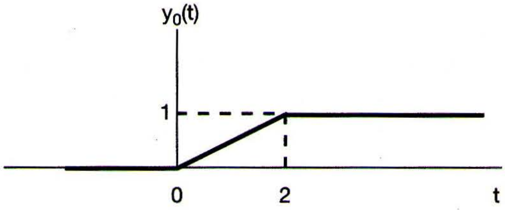


Figure P2.47

In each of these cases, determine whether or not we have enough information to determine the output $y(t)$ when the input is $x(t)$ and the system has impulse response $h(t)$. If it is possible to determine $y(t)$, provide an accurate sketch of it with numerical values clearly indicated on the graph.

2.48. Determine whether each of the following statements concerning LTI systems is true or false. Justify your answers.

(a) If $h(t)$ is the impulse response of an LTI system and $h(t)$ is periodic and nonzero, the system is unstable.

(b) The inverse of a causal LTI system is always causal.

(c) If $|h[n]| \leq K$ for each n, where K is a given number, then the LTI system with $h[n]$ as its impulse response is stable.

(d) If a discrete-time LTI system has an impulse response $h[n]$ of finite duration, the system is stable.

(e) If an LTI system is causal, it is stable.

(f) The cascade of a noncausal LTI system with a causal one is necessarily noncausal.

(g) A continuous-time LTI system is stable if and only if its step response $s(t)$ is absolutely integrable—that is, if and only if

$$\int_{-\infty}^{+\infty} |s(t)|\, dt < \infty.$$

(h) A discrete-time LTI system is causal if and only if its step response $s[n]$ is zero for $n < 0$.

2.49. In the text, we showed that if $h[n]$ is absolutely summable, i.e., if

$$\sum_{k=-\infty}^{+\infty} |h[k]| < \infty,$$

then the LTI system with impulse response $h[n]$ is stable. This means that absolute summability is a *sufficient* condition for stability. In this problem, we shall show that it is also a *necessary* condition. Consider an LTI system with impulse response $h[n]$ that is not absolutely summable; that is,

$$\sum_{k=-\infty}^{+\infty} |h[k]| = \infty.$$

(a) Suppose that the input to this system is

$$x[n] = \begin{cases} 0, & \text{if } h[-n] = 0 \\ \frac{h[-n]}{|h[-n]|}, & \text{if } h[-n] \neq 0 \end{cases}.$$

Does this input signal represent a bounded input? If so, what is the smallest number B such that

$$|x[n]| \leq B \text{ for all } n?$$

(b) Calculate the output at $n = 0$ for this particular choice of input. Does the result prove the contention that absolute summability is a necessary condition for stability?

(c) In a similar fashion, show that a continuous-time LTI system is stable if and only if its impulse response is absolutely integrable.

2.50. Consider the cascade of two systems shown in Figure P2.50. The first system, A, is known to be LTI. The second system, B, is known to be the inverse of system A. Let $y_1(t)$ denote the response of system A to $x_1(t)$, and let $y_2(t)$ denote the response of system A to $x_2(t)$.

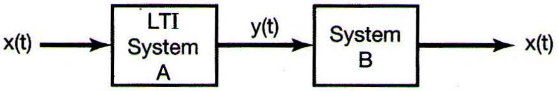


Figure P2.50

(a) What is the response of system B to the input $ay_1(t) + by_2(t)$, where a and b are constants?

(b) What is the response of system B to the input $y_1(t - \tau)$?

2.51. In the text, we saw that the overall input-output relationship of the cascade of two LTI systems does not depend on the order in which they are cascaded. This fact, known as the commutativity property, depends on both the linearity and the time invariance of both systems. In this problem, we illustrate the point.

(a) Consider two discrete-time systems A and B, where system A is an LTI system with unit sample response $h[n] = (1/2)^n u[n]$. System B, on the other hand, is linear but time varying. Specifically, if the input to system B is $w[n]$, its output is

$$z[n] = nw[n].$$

Show that the commutativity property does not hold for these two systems by computing the impulse responses of the cascade combinations in Figures P2.51(a) and P2.51(b), respectively.

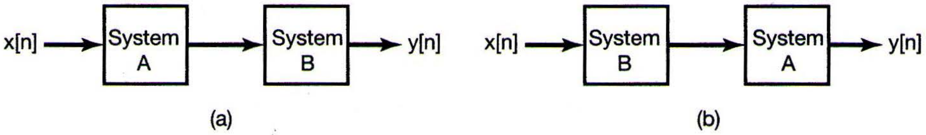


Figure P2.51

(b) Suppose that we replace system B in each of the interconnected systems of Figure P2.51 by the system with the following relationship between its input $w[n]$ and output $z[n]$:

$$z[n] = w[n] + 2.$$

Repeat the calculations of part (a) in this case.

2.52. Consider a discrete-time LTI system with unit sample response

$$h[n] = (n+1)\alpha^n u[n],$$

where $|\alpha| < 1$. Show that the step response of this system is

$$s[n] = \left[\frac{1}{(\alpha-1)^2} - \frac{\alpha}{(\alpha-1)^2}\alpha^n + \frac{\alpha}{(\alpha-1)}(n+1)\alpha^n\right]u[n].$$

(*Hint:* Note that

$$\sum_{k=0}^{N}(k+1)\alpha^k = \frac{d}{d\alpha}\sum_{k=0}^{N+1}\alpha^k.)$$

2.53. **(a)** Consider the homogeneous differential equation

$$\sum_{k=0}^{N} a_k \frac{d^k y(t)}{dt^k} = 0. \qquad \text{(P2.53–1)}$$

Show that if s_0 is a solution of the equation

$$p(s) = \sum_{k=0}^{N} a_k s^k = 0, \qquad \text{(P2.53–2)}$$

then $Ae^{s_0 t}$ is a solution of eq. (P2.53–1), where A is an arbitrary complex constant.

(b) The polynomial $p(s)$ in eq. (P2.53–2) can be factored in terms of its roots $s_1, \ldots, s_r$ as

$$p(s) = a_N(s-s_1)^{\sigma_1}(s-s_2)^{\sigma_2}\ldots(s-s_r)^{\sigma_r},$$

where the s_i are the distinct solutions of eq. (P2.53–2) and the σ_i are their *multiplicities*—that is, the number of times each root appears as a solution of the equation. Note that

$$\sigma_1 + \sigma_2 + \ldots + \sigma_r = N.$$

In general, if $\sigma_i > 1$, then not only is $Ae^{s_i t}$ a solution of eq. (P2.53–1), but so is $At^j e^{s_i t}$, as long as j is an integer greater than or equal to zero and less than or equal to $\sigma_i - 1$. To illustrate this, show that if $\sigma_i = 2$, then $Ate^{s_i t}$ is a solution of eq. (P2.53–1). [*Hint:* Show that if s is an arbitrary complex number, then

$$\sum_{k=0}^{N} \frac{d^k(Ate^{st})}{dt^k} = Ap(s)te^{st} + A\frac{dp(s)}{ds}e^{st}.]$$

Thus, the most general solution of eq. (P2.53–1) is

$$\sum_{i=1}^{r}\sum_{j=0}^{\sigma_i-1} A_{ij} t^j e^{s_i t},$$

where the A_{ij} are arbitrary complex constants.

(c) Solve the following homogeneous differential equations with the specified auxiliary conditions:

(i) $\frac{d^2y(t)}{dt^2} + 3\frac{dy(t)}{dt} + 2y(t) = 0,\ y(0) = 0,\ \ y'(0) = 2$

(ii) $\frac{d^2y(t)}{dt^2} + 3\frac{dy(t)}{dt} + 2y(t) = 0,\ y(0) = 1,\ \ y'(0) = -1$

(iii) $\frac{d^2y(t)}{dt^2} + 3\frac{dy(t)}{dt} + 2y(t) = 0,\ y(0) = 0,\ \ y'(0) = 0$

(iv) $\frac{d^2y(t)}{dt^2} + 2\frac{dy(t)}{dt} + y(t) = 0,\ y(0) = 1,\ \ y'(0) = 1$

(v) $\frac{d^3y(t)}{dt^3} + \frac{d^2y(t)}{dt^2} - \frac{dy(t)}{dt} - y(t) = 0,\ y(0) = 1,\ \ y'(0) = 1,\ \ y''(0) = -2$

(vi) $\frac{d^2y(t)}{dt^2} + 2\frac{dy(t)}{dt} + 5y(t) = 0,\ y(0) = 1,\ \ y'(0) = 1$

2.54. (a) Consider the homogeneous difference equation

$$\sum_{k=0}^{N} a_k y[n-k] = 0, \tag{P2.54–1}$$

Show that if z_0 is a solution of the equation

$$\sum_{k=0}^{N} a_k z^{-k} = 0, \tag{P2.54–2}$$

then Az_0^n is a solution of eq. (P2.54–1), where A is an arbitrary constant.

(b) As it is more convenient for the moment to work with polynomials that have only nonnegative powers of z, consider the equation obtained by multiplying both sides of eq. (P2.54–2) by z^N:

$$p(z) = \sum_{k=0}^{N} a_k z^{N-k} = 0. \tag{P2.54–3}$$

The polynomial $p(z)$ can be factored as

$$p(z) = a_0(z - z_1)^{\sigma_1} \ldots (z - z_r)^{\sigma_r},$$

where the $z_1, \ldots, z_r$ are the distinct roots of $p(z)$.

Show that if $y[n] = nz^{n-1}$, then

$$\sum_{k=0}^{N} a_k y[n-k] = \frac{dp(z)}{dz} z^{n-N} + (n-N)p(z)z^{n-N-1}.$$

Use this fact to show that if $\sigma_i = 2$, then both Az_i^n and Bnz_i^{n-1} are solutions of eq. (P2.54–1), where A and B are arbitrary complex constants. More generally, one can use this same procedure to show that if $\sigma_i > 1$, then

$$A\frac{n!}{r!(n-r)!}z^{n-r}$$

is a solution of eq. (P2.54–1) for $r = 0, 1, \ldots, \sigma_i - 1$.[7]

(c) Solve the following homogeneous difference equations with the specified auxiliary conditions:

(i) $y[n] + \frac{3}{4}y[n-1] + \frac{1}{8}y[n-2] = 0;\ y[0] = 1,\ y[-1] = -6$

(ii) $y[n] - 2y[n-1] + y[n-2] = 0;\ y[0] = 1,\ y[1] = 0$

(iii) $y[n] - 2y[n-1] + y[n-2] = 0;\ y[0] = 1,\ y[10] = 21$

(iv) $y[n] - \frac{\sqrt{2}}{2}y[n-1] + \frac{1}{4}y[n-2] = 0;\ y[0] = 0,\ y[-1] = 1$

2.55. In the text we described one method for solving linear constant-coefficient difference equations, and another method for doing this was illustrated in Problem 2.30. If the assumption of initial rest is made so that the system described by the difference equation is LTI and causal, then, in principle, we can determine the unit impulse response $h[n]$ using either of these procedures. In Chapter 5, we describe another method that allows us to determine $h[n]$ in a more elegant way. In this problem we describe yet another approach, which basically shows that $h[n]$ can be determined by solving the homogeneous equation with appropriate initial conditions.

(a) Consider the system initially at rest and described by the equation

$$y[n] - \frac{1}{2}y[n-1] = x[n]. \qquad \text{(P2.55–1)}$$

Assuming that $x[n] = \delta[n]$, what is $y[0]$? What equation does $h[n]$ satisfy for $n \geq 1$, and with what auxiliary condition? Solve this equation to obtain a closed-form expression for $h[n]$.

(b) Consider next the LTI system initially at rest and described by the difference equation

$$y[n] - \frac{1}{2}y[n-1] = x[n] + 2x[n-1]. \qquad \text{(P2.55–2)}$$

This system is depicted in Figure P2.55(a) as a cascade of two LTI systems that are initially at rest. Because of the properties of LTI systems, we can reverse the order of the systems in the cascade to obtain an alternative representation of the same overall system, as illustrated in Figure P2.55(b). From this fact, use the result of part (a) to determine the impulse response for the system described by eq. (P2.55–2).

(c) Consider again the system of part (a), with $h[n]$ denoting its impulse response. Show, by verifying that eq. (P2.55–3) satisfies the difference equation (P2.55–1), that the response $y[n]$ to an arbitrary input $x[n]$ is in fact given by the convolution sum

$$y[n] = \sum_{m=-\infty}^{+\infty} h[n-m]x[m]. \qquad \text{(P2.55–3)}$$

---

[7] Here, we are using factorial notation—that is, $k! = k(k-1)(k-2)\ldots(2)(1)$, where $0!$ is defined to be 1.

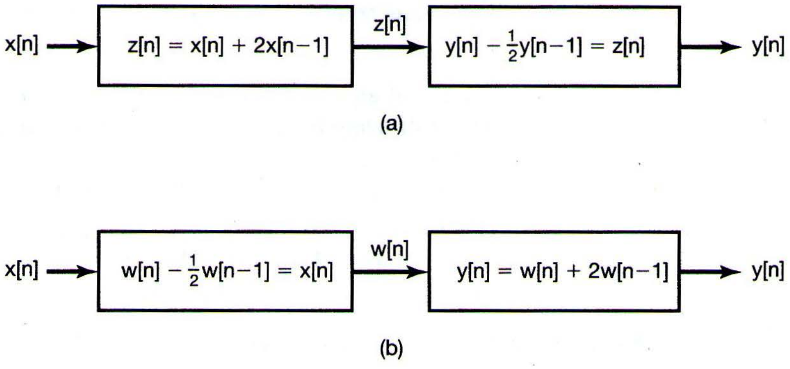


Figure P2.55

(d) Consider the LTI system initially at rest and described by the difference equation

$$\sum_{k=0}^{N} a_k y[n-k] = x[n]. \qquad \text{(P2.55–4)}$$

Assuming that $a_0 \neq 0$, what is $y[0]$ if $x[n] = \delta[n]$? Using this result, specify the homogeneous equation and initial conditions that the impulse response of the system must satisfy.

Consider next the causal LTI system described by the difference equation

$$\sum_{k=0}^{N} a_k y[n-k] = \sum_{k=0}^{M} b_k x[n-k]. \qquad \text{(P2.55–5)}$$

Express the impulse response of this system in terms of that for the LTI system described by eq. (P2.55–4).

(e) There is an alternative method for determining the impulse response of the LTI system described by eq. (P2.55–5). Specifically, given the condition of initial rest, i.e., in this case, $y[-N] = y[-N+1] = \ldots = y[-1] = 0$, solve eq. (P2.55–5) recursively when $x[n] = \delta[n]$ in order to determine $y[0], \ldots, y[M]$. What equation does $h[n]$ satisfy for $n \geq M$? What are the appropriate initial conditions for this equation?

(f) Using either of the methods outlined in parts (d) and (e), find the impulse responses of the causal LTI systems described by the following equations:

(i) $y[n] - y[n-2] = x[n]$

(ii) $y[n] - y[n-2] = x[n] + 2x[n-1]$

(iii) $y[n] - y[n-2] = 2x[n] - 3x[n-4]$

(iv) $y[n] - (\sqrt{3}/2)y[n-1] + \frac{1}{4}y[n-2] = x[n]$

2.56. In this problem, we consider a procedure that is the continuous-time counterpart of the technique developed in Problem 2.55. Again, we will see that the problem of determining the impulse response $h(t)$ for $t > 0$ for an LTI system initially at rest and described by a linear constant-coefficient differential equation reduces to the problem of solving the homogeneous equation with appropriate initial conditions.

(a) Consider the LTI system initially at rest and described by the differential equation

$$\frac{dy(t)}{dt} + 2y(t) = x(t). \tag{P2.56–1}$$

Suppose that $x(t) = \delta(t)$. In order to determine the value of $y(t)$ *immediately after* the application of the unit impulse, consider integrating eq. (P2.56–1) from $t = 0^-$ to $t = 0^+$ (i.e., from "just before" to "just after" the application of the impulse). This yields

$$y(0^+) - y(0^-) + 2\int_{0^-}^{0^+} y(\tau)\,d\tau = \int_{0^-}^{0^+} \delta(\tau)\,d\tau = 1. \tag{P2.56–2}$$

Since the system is initially at rest and $x(t) = 0$ for $t < 0$, $y(0^-) = 0$. To satisfy eq. (P2.56–2) we must have $y(0^+) = 1$. Thus, since $x(t) = 0$ for $t > 0$, the impulse response of our system is the solution of the homogeneous differential equation

$$\frac{dy(t)}{dt} + 2y(t) = 0$$

with initial condition

$$y(0^+) = 1.$$

Solve this differential equation to obtain the impulse response $h(t)$ for the system. Check your result by showing that

$$y(t) = \int_{-\infty}^{+\infty} h(t-\tau)x(\tau)\,d\tau$$

satisfies eq. (P2.56–1) for any input $x(t)$.

(b) To generalize the preceding argument, consider an LTI system initially at rest and described by the differential equation

$$\sum_{k=0}^{N} a_k \frac{d^k y(t)}{dt^k} = x(t) \tag{P2.56–3}$$

with $x(t) = \delta(t)$. Assume the condition of initial rest, which, since $x(t) = 0$ for $t < 0$, implies that

$$y(0^-) = \frac{dy}{dt}(0^-) = \ldots = \frac{d^{N-1}y}{dt^{N-1}}(0^-) = 0. \tag{P2.56–4}$$

Integrate both sides of eq. (P2.56–3) once from $t = 0^-$ to $t = 0^+$, and use eq. (P2.56–4) and an argument similar to that used in part (a) to show that the resulting equation is satisfied with

$$y(0^+) = \frac{dy}{dt}(0^+) = \ldots = \frac{d^{N-2}y}{dt^{N-2}}(0^+) = 0 \qquad \text{(P2.56–5}a\text{)}$$

and

$$\frac{d^{N-1}y}{dt^{N-1}}(0^+) = \frac{1}{a_N}. \qquad \text{(P2.56–5}b\text{)}$$

Consequently, the system's impulse response for $t > 0$ can be obtained by solving the homogeneous equation

$$\sum_{k=0}^{N} a_k \frac{d^k y(t)}{dt^k} = 0$$

with initial conditions given by eqs. (P2.56–5).

(c) Consider now the causal LTI system described by the differential equation

$$\sum_{k=0}^{N} a_k \frac{d^k y(t)}{dt^k} = \sum_{k=0}^{M} b_k \frac{d^k x(t)}{dt^k}. \qquad \text{(P2.56–6)}$$

Express the impulse response of this system in terms of that for the system of part (b). (*Hint:* Examine Figure P2.56.)

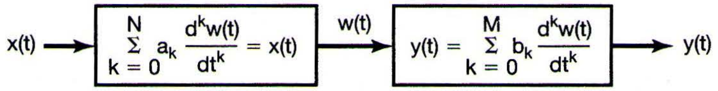

x(t)
$\sum_{k=0}^{N} a_k \frac{d^k w(t)}{dt^k} = x(t)$
w(t)
$y(t) = \sum_{k=0}^{M} b_k \frac{d^k w(t)}{dt^k}$
y(t)

Figure P2.56

(d) Apply the procedures outlined in parts (b) and (c) to find the impulse responses for the LTI systems initially at rest and described by the following differential equations:

(i) $\frac{d^2y(t)}{dt^2} + 3\frac{dy(t)}{dt} + 2y(t) = x(t)$

(ii) $\frac{d^2y(t)}{dt^2} + 2\frac{dy(t)}{dt} + 2y(t) = x(t)$

(e) Use the results of parts (b) and (c) to deduce that if $M \geq N$ in eq. (P2.56–6), then the impulse response $h(t)$ will contain singularity terms concentrated at $t = 0$. In particular, $h(t)$ will contain a term of the form

$$\sum_{r=0}^{M-N} \alpha_r u_r(t),$$

where the α_r are constants and the $u_r(t)$ are the singularity functions defined in Section 2.5.

(f) Find the impulse responses of the causal LTI systems described by the following differential equations:

(i) $\frac{dy(t)}{dt} + 2y(t) = 3\frac{dx(t)}{dt} + x(t)$

(ii) $\frac{d^2y(t)}{dt^2} + 5\frac{dy(t)}{dt} + 6y(t) = \frac{d^3x(t)}{dt^3} + 2\frac{d^2x(t)}{dt^2} + 4\frac{dx(t)}{dt} + 3x(t)$

2.57. Consider a causal LTI system S whose input $x[n]$ and output $y[n]$ are related by the difference equation

$$y[n] = -ay[n-1] + b_0x[n] + b_1x[n-1].$$

(a) Verify that S may be considered a cascade connection of two causal LTI systems S_1 and S_2 with the following input-output relationship:

$$\begin{aligned} S_1 : y_1[n] &= b_0x_1[n] + b_1x_1[n-1], \\ S_2 : y_2[n] &= -ay_2[n-1] + x_2[n]. \end{aligned}$$

(b) Draw a block diagram representation of S_1.
(c) Draw a block diagram representation of S_2.
(d) Draw a block diagram representation of S as a cascade connection of the block diagram representation of S_1 followed by the block diagram representation of S_2.
(e) Draw a block diagram representation of S as a cascade connection of the block diagram representation of S_2 followed by the block diagram representation of S_1.
(f) Show that the two unit-delay elements in the block diagram representation of S obtained in part (e) may be collapsed into one unit-delay element. The resulting block diagram is referred to as a *Direct Form II* realization of S, while the block diagrams obtained in parts (d) and (e) are referred to as *Direct Form I* realizations of S.

2.58. Consider a causal LTI system S whose input $x[n]$ and output $y[n]$ are related by the difference equation

$$2y[n] - y[n-1] + y[n-3] = x[n] - 5x[n-4].$$

(a) Verify that S may be considered a cascade connection of two causal LTI systems S_1 and S_2 with the following input-output relationship:

$$\begin{aligned} S_1 : 2y_1[n] &= x_1[n] - 5x_1[n-4], \\ S_2 : y_2[n] &= \frac{1}{2}y_2[n-1] - \frac{1}{2}y_2[n-3] + x_2[n]. \end{aligned}$$

(b) Draw a block diagram representation of S_1.
(c) Draw a block diagram representation of S_2.
(d) Draw a block diagram representation of S as a cascade connection of the block diagram representation of S_1 followed by the block diagram representation of S_2.
(e) Draw a block diagram representation of S as a cascade connection of the block diagram representation of S_2 followed by the block diagram representation of S_1.
(f) Show that the four delay elements in the block diagram representation of S obtained in part (e) may be collapsed to three. The resulting block diagram is referred to as a *Direct Form II* realization of S, while the block diagrams obtained in parts (d) and (e) are referred to as *Direct Form I* realizations of S.

2.59. Consider a causal LTI system S whose input $x(t)$ and output $y(t)$ are related by the differential equation

$$a_1 \frac{dy(t)}{dt} + a_0 y(t) = b_0 x(t) + b_1 \frac{dx(t)}{dt}.$$

(a) Show that

$$y(t) = A\int_{-\infty}^{t} y(\tau)\,d\tau + Bx(t) + C\int_{-\infty}^{t} x(\tau)\,d\tau,$$

and express the constants A, B, and C in terms of the constants a_0, a_1, b_0, and b_1.

(b) Show that S may be considered a cascade connection of the following two causal LTI systems:

$$S_1 : y_1(t) = Bx_1(t) + C\int_{-\infty}^{t} x(\tau)\,d\tau,$$

$$S_2 : y_2(t) = A\int_{-\infty}^{t} y_2(\tau)\,d\tau + x_2(t).$$

(c) Draw a block diagram representation of S_1.

(d) Draw a block diagram representation of S_2.

(e) Draw a block diagram representation of S as a cascade connection of the block diagram representation of S_1 followed by the block diagram representation of S_2.

(f) Draw a block diagram representation of S as a cascade connection of the block diagram representation of S_2 followed by the block diagram of representation S_1.

(g) Show that the two integrators in your answer to part (f) may be collapsed into one. The resulting block diagram is referred to as a *Direct Form II* realization of S, while the block diagrams obtained in parts (e) and (f) are referred to as *Direct Form I* realizations of S.

2.60. Consider a causal LTI system S whose input $x(t)$ and output $y(t)$ are related by the differential equation

$$a_2 \frac{d^2y(t)}{dt^2} + a_1 \frac{dy(t)}{dt} + a_0 y(t) = b_0 x(t) + b_1 \frac{dx(t)}{dt} + b_2 \frac{d^2x(t)}{dt^2}.$$

(a) Show that

$$y(t) = A\int_{-\infty}^{t} y(\tau)\,d\tau + B\int_{-\infty}^{t}\left(\int_{-\infty}^{\tau} y(\sigma)\,d\sigma\right)d\tau$$
$$+ Cx(t) + D\int_{-\infty}^{t} x(\tau)\,d\tau + E\int_{-\infty}^{t}\left(\int_{-\infty}^{\tau} x(\sigma)\,d\sigma\right)d\tau,$$

and express the constants A, B, C, D, and E in terms of the constants a_0, a_1, a_2, b_0, b_1, and b_2.

(b) Show that S may be considered a cascade connection of the following two causal LTI systems:

$$S_1 : y_1(t) = Cx_1(t) + D\int_{-\infty}^{t} x_1(\tau)\,d\tau + E\int_{-\infty}^{t}\left(\int_{-\infty}^{\tau} x_1(\sigma)\,d\sigma\right)d\tau,$$

$$S_2 : y_2(t) = A\int_{-\infty}^{t} y_2(\tau)\,d\tau + B\int_{-\infty}^{t}\left(\int_{-\infty}^{\tau} y_2(\sigma)\,d\sigma\right)d\tau + x_2(t).$$

(c) Draw a block diagram representation of S_1.

(d) Draw a block diagram representation of S_2.

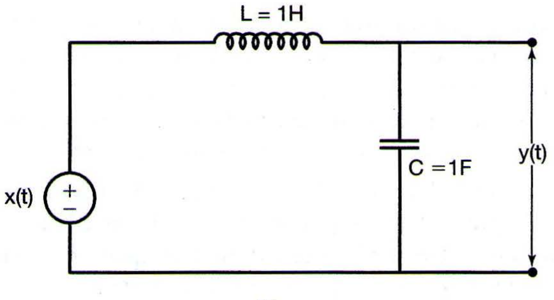
(e) Draw a block diagram representation of S as a cascade connection of the block diagram representation of S_1 followed by the block diagram representation of S_2.

(f) Draw a block diagram representation of S as a cascade connection of the block diagram representation of S_2 followed by the block diagram representation of S_1.

(g) Show that the four integrators in your answer to part (f) may be collapsed into two. The resulting block diagram is referred to as a *Direct Form II* realization of S, while the block diagrams obtained in parts (e) and (f) are referred to as *Direct Form I* realizations of S.

EXTENSION PROBLEMS

2.61. **(a)** In the circuit shown in Figure P2.61(a), $x(t)$ is the input voltage. The voltage $y(t)$ across the capacitor is considered to be the system output.



(a)

Figure P2.61a

(i) Determine the differential equation relating $x(t)$ and $y(t)$.

(ii) Show that the homogeneous solution of the differential equation from part (i) has the form $K_1e^{j\omega_1 t} + K_2e^{j\omega_2 t}$. Specify the values of ω_1 and ω_2.

(iii) Show that, since the voltage and current are restricted to be real, the natural response of the system is sinusoidal.

(b) In the circuit shown in Figure P2.61(b), $x(t)$ is the input voltage. The voltage $y(t)$ across the capacitor is considered to be the system output.

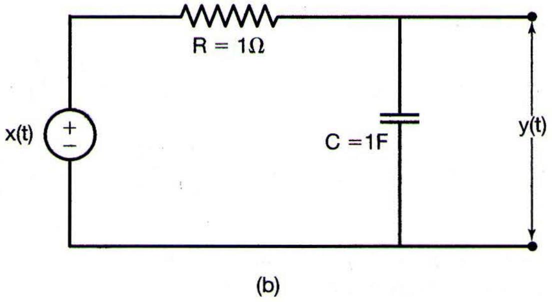


Figure P2.61b

(i) Determine the differential equation relating $x(t)$ and $y(t)$.

(ii) Show that the natural response of this system has the form Ke^{-at}, and specify the value of a.

(c) In the circuit shown in Figure P2.61(c), $x(t)$ is the input voltage. The voltage $y(t)$ across the capacitor is considered to be the system output.

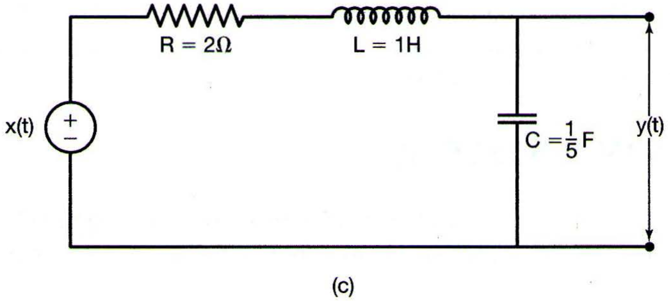


Figure P2.61c

(i) Determine the differential equation relating $x(t)$ and $y(t)$.

(ii) Show that the homogeneous solution of the differential equation from part (i) has the form $e^{-at}\{K_1e^{j2t} + K_2e^{-j2t}\}$, and specify the value of a.

(iii) Show that, since the voltage and current are restricted to be real, the natural response of the system is a decaying sinusoid.

2.62. **(a)** In the mechanical system shown in Figure P2.62(a), the force $x(t)$ applied to the mass represents the input, while the displacement $y(t)$ of the mass represents the output. Determine the differential equation relating $x(t)$ and $y(t)$. Show that the natural response of this system is periodic.

(b) Consider Figure P2.62(b), in which the force $x(t)$ is the input and the velocity $y(t)$ is the output. The mass of the car is m, while the coefficient of kinetic friction is ρ. Show that the natural response of this system decays with increasing time.

(c) In the mechanical system shown in Figure P2.62(c), the force $x(t)$ applied to the mass represents the input, while the displacement $y(t)$ of the mass represents the output.

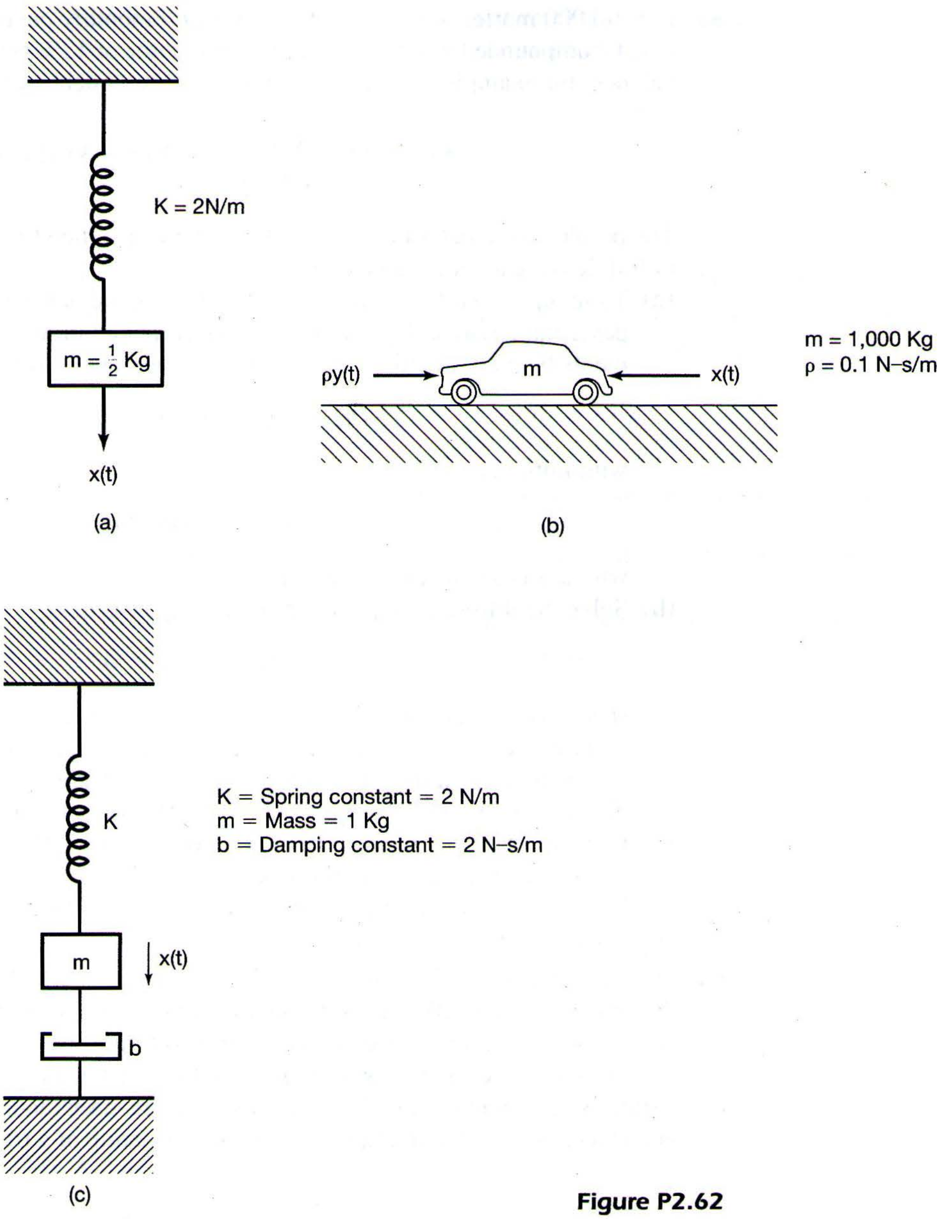

K = 2N/m
m = $\frac{1}{2}$ Kg
x(t)
(a)
m = 1,000 Kg
ρ = 0.1 N–s/m
ρy(t)
m
x(t)
(b)
K = Spring constant = 2 N/m
m = Mass = 1 Kg
b = Damping constant = 2 N–s/m
K
m
x(t)
b
(c)

Figure P2.62

(i) Determine the differential equation relating $x(t)$ and $y(t)$.

(ii) Show that the homogeneous solution of the differential equation from part (i) has the form $e^{-at}\{K_1 e^{jt} + K_2 e^{-jt}\}$, and specify the value of a.

(iii) Show that, since the force and displacement are restricted to be real, the natural response of the system is a decaying sinusoid.

2.63. A \$100,000 mortgage is to be retired by *equal* monthly payments of D dollars. Interest, compounded monthly, is charged at the rate of 12% per annum on the unpaid balance; for example, after the first month, the total debt equals

$$\$100{,}000 + \left(\frac{0.12}{12}\right)\$100{,}000 = \$101{,}000.$$

The problem is to determine D such that after a specified time the mortgage is paid in full, leaving a net balance of zero.

(a) To set up the problem, let $y[n]$ denote the unpaid balance after the nth monthly payment. Assume that the principal is borrowed in month 0 and monthly payments begin in month 1. Show that $y[n]$ satisfies the difference equation

$$y[n] - \gamma y[n-1] = -D \quad n \geq 1 \tag{P2.63–1}$$

with initial condition

$$y[0] = \$100{,}000,$$

where γ is a constant. Determine γ.

(b) Solve the difference equation of part (a) to determine

$$y[n] \quad \text{for } n \geq 0.$$

(*Hint:* The particular solution of eq. (P2.63–1) is a constant Y. Find the value of Y, and express $y[n]$ for $n \geq 1$ as the sum of particular and homogeneous solutions. Determine the unknown constant in the homogeneous solution by directly calculating $y[1]$ from eq. (P2.63–1) and comparing it to your solution.)

(c) If the mortgage is to be retired in 30 years after 360 monthly payments of D dollars, determine the appropriate value of D.

(d) What is the total payment to the bank over the 30-year period?

(e) Why do banks make loans?

2.64. One important use of inverse systems is in situations in which one wishes to remove distortions of some type. A good example of this is the problem of removing echoes from acoustic signals. For example, if an auditorium has a perceptible echo, then an initial acoustic impulse will be followed by attenuated versions of the sound at regularly spaced intervals. Consequently, an often-used model for this phenomenon is an LTI system with an impulse response consisting of a train of impulses, i.e.,

$$h(t) = \sum_{k=0}^{\infty} h_k \delta(t - kT). \tag{P2.64–1}$$

Here the echoes occur T seconds apart, and h_k represents the gain factor on the kth echo resulting from an initial acoustic impulse.

(a) Suppose that $x(t)$ represents the original acoustic signal (the music produced by an orchestra, for example) and that $y(t) = x(t) * h(t)$ is the actual signal that is heard if no processing is done to remove the echoes. In order to remove the distortion introduced by the echoes, assume that a microphone is used to sense $y(t)$ and that the resulting signal is transduced into an electrical signal. We will also use $y(t)$ to denote this signal, as it represents the electrical equivalent of the acoustic signal, and we can go from one to the other via acoustic-electrical conversion systems.

The important point to note is that the system with impulse response given by eq. (P2.64–1) is invertible. Therefore, we can find an LTI system with impulse response $g(t)$ such that

$$y(t) * g(t) = x(t),$$

and thus, by processing the electrical signal $y(t)$ in this fashion and then converting back to an acoustic signal, we can remove the troublesome echoes.

The required impulse response $g(t)$ is also an impulse train:

$$g(t) = \sum_{k=0}^{\infty} g_k \delta(t - kT).$$

Determine the algebraic equations that the successive g_k must satisfy, and solve these equations for g_0, g_1, and g_2 in terms of h_k.

(b) Suppose that $h_0 = 1$, $h_1 = 1/2$, and $h_i = 0$ for all $i \geq 2$. What is $g(t)$ in this case?

(c) A good model for the generation of echoes is illustrated in Figure P2.64. Hence, each successive echo represents a fed-back version of $y(t)$, delayed by T seconds and scaled by α. Typically, $0 < \alpha < 1$, as successive echoes are attenuated.

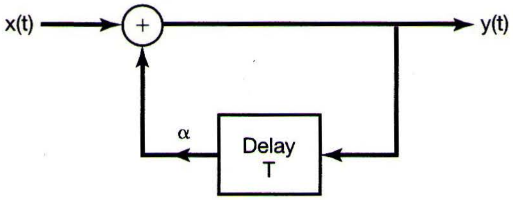


Figure P2.64

(i) What is the impulse response of this system? (Assume initial rest, i.e., $y(t) = 0$ for $t < 0$ if $x(t) = 0$ for $t < 0$.)

(ii) Show that the system is stable if $0 < \alpha < 1$ and unstable if $\alpha > 1$.

(iii) What is $g(t)$ in this case? Construct a realization of the inverse system using adders, coefficient multipliers, and T-second delay elements.

(d) Although we have phrased the preceding discussion in terms of continuous-time systems because of the application we have been considering, the same general ideas hold in discrete time. That is, the LTI system with impulse response

$$h[n] = \sum_{k=0}^{\infty} h_k \delta[n - kN]$$

is invertible and has as its inverse an LTI system with impulse response

$$g[n] = \sum_{k=0}^{\infty} g_k \delta[n - kN].$$

It is not difficult to check that the g_k satisfy the same algebraic equations as in part (a).

Consider now the discrete-time LTI system with impulse response

$$h[n] = \sum_{k=-\infty}^{\infty} \delta[n - kN].$$

This system is *not* invertible. Find two inputs that produce the same output.

2.65. In Problem 1.45, we introduced and examined some of the basic properties of correlation functions for continuous-time signals. The discrete-time counterpart of the correlation function has essentially the same properties as those in continuous time, and both are extremely important in numerous applications (as is discussed in Problems 2.66 and 2.67). In this problem, we introduce the discrete-time correlation function and examine several more of its properties.

Let $x[n]$ and $y[n]$ be two real-valued discrete-time signals. The *autocorrelation functions* $\phi_{xx}[n]$ and $\phi_{yy}[n]$ of $x[n]$ and $y[n]$, respectively, are defined by the expressions

$$\phi_{xx}[n] = \sum_{m=-\infty}^{+\infty} x[m+n]x[m]$$

and

$$\phi_{yy}[n] = \sum_{m=-\infty}^{+\infty} y[m+n]y[m],$$

and the *cross-correlation functions* are given by

$$\phi_{xy}[n] = \sum_{m=-\infty}^{+\infty} x[m+n]y[m]$$

and

$$\phi_{yx}[n] = \sum_{m=-\infty}^{+\infty} y[m+n]x[m].$$

As in continuous time, these functions possess certain symmetry properties. Specifically, $\phi_{xx}[n]$ and $\phi_{yy}[n]$ are even functions, while $\phi_{xy}[n] = \phi_{yx}[-n]$.

(a) Compute the autocorrelation sequences for the signals $x_1[n]$, $x_2[n]$, $x_3[n]$, and $x_4[n]$ depicted in Figure P2.65.

(b) Compute the cross-correlation sequences

$$\phi_{x_i x_j}[n], \ i \neq j, \ \ i, j = 1, 2, 3, 4,$$

for $x_i[n]$, $i = 1, 2, 3, 4$, as shown in Figure P2.65.

(c) Let $x[n]$ be the input to an LTI system with unit sample response $h[n]$, and let the corresponding output be $y[n]$. Find expressions for $\phi_{xy}[n]$ and $\phi_{yy}[n]$ in terms of $\phi_{xx}[n]$ and $h[n]$. Show how $\phi_{xy}[n]$ and $\phi_{yy}[n]$ can be viewed as the output

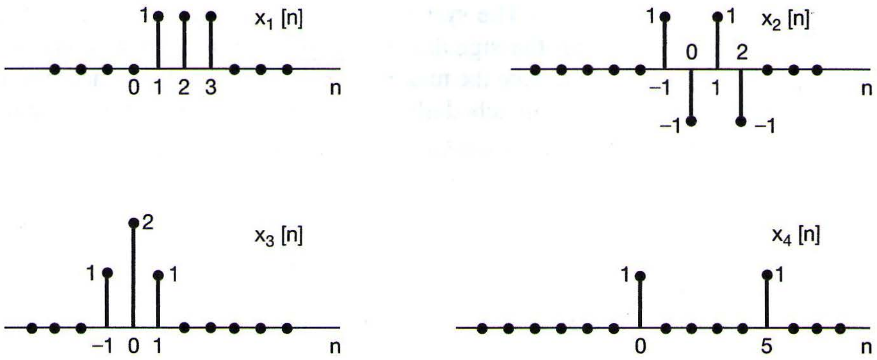


Figure P2.65

of LTI systems with $\phi_{xx}[n]$ as the input. (Do this by explicitly specifying the impulse response of each of the two systems.)

(d) Let $h[n] = x_1[n]$ in Figure P2.65, and let $y[n]$ be the output of the LTI system with impulse response $h[n]$ when the input $x[n]$ also equals $x_1[n]$. Calculate $\phi_{xy}[n]$ and $\phi_{yy}[n]$ using the results of part (c).

2.66. Let $h_1(t)$, $h_2(t)$, and $h_3(t)$, as sketched in Figure P2.66, be the impulse responses of three LTI systems. These three signals are known as *Walsh functions* and are of considerable practical importance because they can be easily generated by digital logic circuitry and because multiplication by each of them can be implemented in a simple fashion by a polarity-reversing switch.

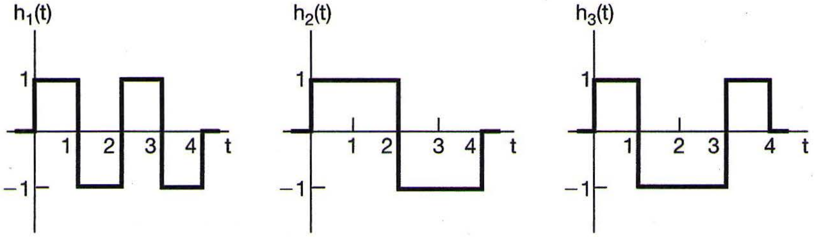


Figure P2.66

(a) Determine and sketch a choice for $x_1(t)$, a continuous-time signal with the following properties:

(i) $x_1(t)$ is real.

(ii) $x_1(t) = 0$ for $t < 0$.

(iii) $|x_1(t)| \leq 1$ for $t \geq 0$.

(iv) $y_1(t) = x_1(t) * h(t)$ is as large as possible at $t = 4$.

(b) Repeat part (a) for $x_2(t)$ and $x_3(t)$ by making $y_2(t) = x_2(t) * h_2(t)$ and $y_3(t) = x_3(t) * h_3(t)$ each as large as possible at $t = 4$.

(c) What is the value of

$$y_{ij}(t) = x_i(t) * h_j(t), \; i \neq j$$

at time $t = 4$ for $i, j = 1, 2, 3$?

The system with impulse response $h_i(t)$ is known as the *matched filter* for the signal $x_i(t)$ because the impulse response is tuned to $x_i(t)$ in order to produce the maximum output signal. In the next problem, we relate the concept of a matched filter to that of the correlation function for continuous-time signals.

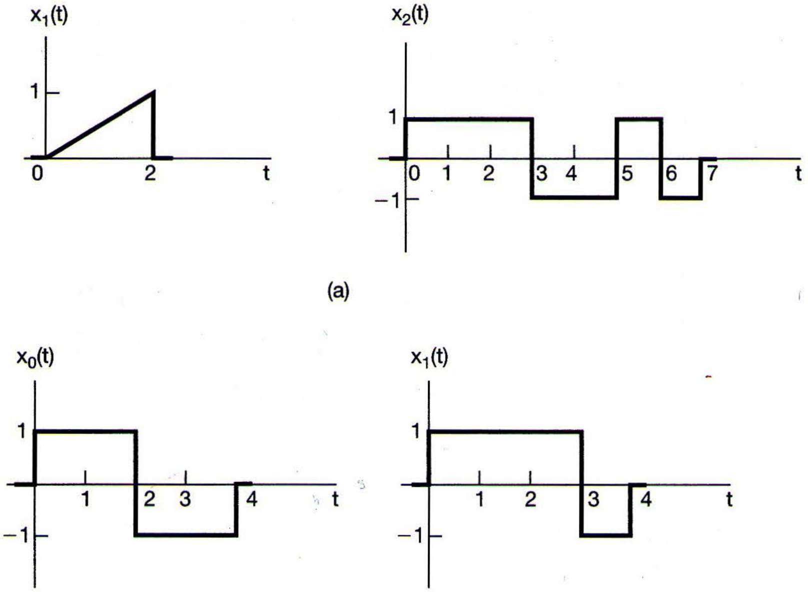
2.67. The *cross-correlation function* between two continuous-time real signals $x(t)$ and $y(t)$ is

$$\phi_{xy}(t) = \int_{-\infty}^{+\infty} x(t+\tau)y(\tau)\,d\tau. \tag{P2.67–1}$$

The *autocorrelation function* of a signal $x(t)$ is obtained by setting $y(t) = x(t)$ in eq. (P2.67–1):

$$\phi_{xx}(t) = \int_{-\infty}^{+\infty} x(t+\tau)x(\tau)\,d\tau.$$

(a) Compute the autocorrelation function for each of the two signals $x_1(t)$ and $x_2(t)$ depicted in Figure P2.67(a).



(a)

(b)

Figure P2.67

(b) Let $x(t)$ be a given signal, and assume that $x(t)$ is of finite duration—i.e., that $x(t) = 0$ for $t < 0$ and $t > T$. Find the impulse response of an LTI system so that $\phi_{xx}(t - T)$ is the output if $x(t)$ is the input.

(c) The system determined in part (b) is a *matched filter* for the signal $x(t)$. That this definition of a matched filter is identical to the one introduced in Problem 2.66 can be seen from the following:

Let $x(t)$ be as in part (b), and let $y(t)$ denote the response to $x(t)$ of an LTI system with real impulse response $h(t)$. Assume that $h(t) = 0$ for $t < 0$ and for $t > T$. Show that the choice for $h(t)$ that maximizes $y(T)$, subject to the constraint that

$$\int_0^T h^2(t)dt = M, \text{ a fixed positive number,} \qquad \text{(P2.67–2)}$$

is a scalar multiple of the impulse response determined in part (b). [*Hint:* Schwartz's inequality states that

$$\int_b^a u(t)v(t)dt \leq \left[\int_a^b u^2(t)dt\right]^{1/2} \left[\int_a^b v^2(t)dt\right]^{1/2}$$

for any two signals $u(t)$ and $v(t)$. Use this to obtain a bound on y(T).]

(d) The constraint given by eq. (P2.67–2) simply provides a scaling to the impulse response, as increasing M merely changes the scalar multiplier mentioned in part (c). Thus, we see that the particular choice for $h(t)$ in parts (b) and (c) is matched to the signal $x(t)$ to produce maximum output. This is an extremely important property in a number of applications, as we will now indicate.

In communication problems, one often wishes to transmit one of a small number of possible pieces of information. For example, if a complex message is encoded into a sequence of binary digits, we can imagine a system that transmits the information bit by bit. Each bit can then be transmitted by sending one signal, say, $x_0(t)$, if the bit is a 0, or a different signal $x_1(t)$ if a 1 is to be communicated. In this case, the receiving system for these signals must be capable of recognizing whether $x_0(t)$ or $x_1(t)$ has been received. Intuitively, what makes sense is to have two systems in the receiver, one tuned to $x_0(t)$ and one tuned to $x_1(t)$, where, by "tuned," we mean that the system gives a large output after the signal to which it is tuned is received. The property of producing a large output when a particular signal is received is exactly what the matched filter possesses.

In practice, there is always distortion and interference in the transmission and reception processes. Consequently, we want to maximize the difference between the response of a matched filter to the input to which it is matched and the response of the filter to one of the other signals that can be transmitted. To illustrate this point, consider the two signals $x_0(t)$ and $x_1(t)$ depicted in Figure P2.67(b). Let L_0 denote the matched filter for $x_0(t)$, and let L_1 denote the matched filter for $x_1(t)$.

(i) Sketch the responses of L_0 to $x_0(t)$ and $x_1(t)$. Do the same for L_1.

(ii) Compare the values of these responses at $t = 4$. How might you modify $x_0(t)$ so that the receiver would have an even easier job of distinguishing between $x_0(t)$ and $x_1(t)$ in that the response of L_0 to $x_1(t)$ and L_1 to $x_0(t)$ would both be zero at $t = 4$?

2.68. Another application in which matched filters and correlation functions play an important role is radar systems. The underlying principle of radar is that an electromagnetic pulse transmitted at a target will be reflected by the target and will subsequently return to the sender with a delay proportional to the distance to the target. Ideally, the received signal will simply be a shifted and possibly scaled version of the original transmitted signal.

Let $p(t)$ be the original pulse that is sent out. Show that

$$\phi_{pp}(0) = \max_t \phi_{pp}(t).$$

That is, $\phi_{pp}(0)$ is the largest value taken by $\phi_{pp}(t)$. Use this equation to deduce that, if the waveform that comes back to the sender is

$$x(t) = \alpha p(t - t_0),$$

where α is a positive constant, then

$$\phi_{xp}(t_0) = \max_t \phi_{xp}(t).$$

(*Hint:* Use Schwartz's inequality.)

Thus, the way in which simple radar ranging systems work is based on using a matched filter for the transmitted waveform $p(t)$ and noting the time at which the output of this system reaches its maximum value.

2.69. In Section 2.5, we characterized the unit doublet through the equation

$$x(t) * u_1(t) = \int_{-\infty}^{+\infty} x(t-\tau)u_1(\tau)\,d\tau = x'(t) \tag{P2.69–1}$$

for any signal $x(t)$. From this equation, we derived the relationship

$$\int_{-\infty}^{+\infty} g(\tau)u_1(\tau)\,d\tau = -g'(0). \tag{P2.69–2}$$

(a) Show that eq. (P2.69–2) is an equivalent characterization of $u_1(t)$ by showing that eq. (P2.69–2) implies eq. (P2.69–1). [*Hint:* Fix t, and define the signal $g(\tau) = x(t-\tau)$.]

Thus, we have seen that characterizing the unit impulse or unit doublet by how it behaves under convolution is equivalent to characterizing how it behaves under integration when multiplied by an arbitrary signal $g(t)$. In fact, as indicated in Section 2.5, the equivalence of these operational definitions holds for all signals and, in particular, for all singularity functions.

(b) Let $f(t)$ be a given signal. Show that

$$f(t)u_1(t) = f(0)u_1(t) - f'(0)\delta(t)$$

by showing that both functions have the same operational definitions.

(c) What is the value of

$$\int_{-\infty}^{\infty} x(\tau)u_2(\tau)\,d\tau?$$

Find an expression for $f(t)u_2(t)$ analogous to that in part (b) for $f(t)u_1(t)$.

2.70. In analogy with continuous-time singularity functions, we can define a set of discrete-time signals. Specifically, let

$$\begin{aligned} u_{-1}[n] &= u[n], \\ u_0[n] &= \delta[n], \end{aligned}$$

and

$$u_1[n] = \delta[n] - \delta[n-1],$$

and define

$$u_k[n] = \underbrace{u_1[n] * u_1[n] * \cdots * u_1[n]}_{k \text{ times}}, \quad k > 0$$

and

$$u_k[n] = \underbrace{u_{-1}[n] * u_{-1}[n] * \cdots * u_{-1}[n]}_{|k| \text{ times}}, \quad k < 0.$$

Note that

$$\begin{aligned} x[n] * \delta[n] &= x[n], \\ x[n] * u[n] &= \sum_{m=-\infty}^{\infty} x[m], \end{aligned}$$

and

$$x[n] * u_1[n] = x[n] - x[n-1],$$

(a) What is

$$\sum_{m=\infty}^{\infty} x[m]u_1[m]?$$

(b) Show that

$$\begin{aligned} x[n]u_1[n] &= x[0]u_1[n] - [x[1] - x[0]]\delta[n-1] \\ &= x[1]u_1[n] - [x[1] - x[0]]\delta[n]. \end{aligned}$$

(c) Sketch the signals $u_2[n]$ and $u_3[n]$.
(d) Sketch $u_{-2}[n]$ and $u_{-3}[n]$.
(e) Show that, in general, for $k > 0$,

$$u_k[n] = \frac{(-1)^n k!}{n!(k-n)!}[u[n] - u[n-k-1]]. \qquad \text{(P2.70–1)}$$

(*Hint:* Use induction. From part (c), it is evident that $u_k[n]$ satisfies eq. (P2.70–1) for $k = 2$ and 3. Then, assuming that eq. (P2.70–1) satisfies $u_k[n]$, write $u_{k+1}[n]$ in terms of $u_k[n]$, and show that the equation also satisfies $u_{k+1}[n]$.)

(f) Show that, in general, for $k > 0$,

$$u_{-k}[n] = \frac{(n+k-1)!}{n!(k-1)!}u[n]. \tag{P2.70–2}$$

(*Hint:* Again, use induction. Note that

$$u_{-(k+1)}[n] - u_{-(k+1)}[n-1] = u_{-k}[n]. \tag{P2.70–3}$$

Then, assuming that eq. (P2.70–2) is valid for $u_{-k}[n]$, use eq. (P2.70–3) to show that eq. (P2.70–2) is valid for $u_{-(k+1)}[n]$ as well.)

2.71. In this chapter, we have used several properties and ideas that greatly facilitate the analysis of LTI systems. Among these are two that we wish to examine a bit more closely. As we will see, in certain very special cases one must be careful in using these properties, which otherwise hold without qualification.

(a) One of the basic and most important properties of convolution (in both continuous and discrete time) is associativity. That is, if $x(t)$, $h(t)$, and $g(t)$ are three signals, then

$$x(t) * [g(t) * h(t)] = [x(t) * g(t)] * h(t) = [x(t) * h(t)] * g(t). \tag{P2.71–1}$$

This relationship holds as long as all three expressions are well defined and finite. As that is usually the case in practice, we will in general use the associativity property without comments or assumptions. However, there are some cases in which it does *not* hold. For example, consider the system depicted in Figure P2.71, with $h(t) = u_1(t)$ and $g(t) = u(t)$. Compute the response of this system to the input

$$x(t) = 1 \text{ for all } t.$$

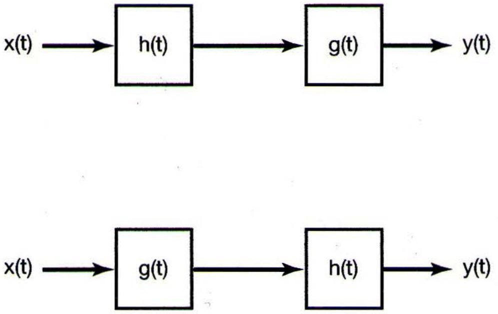


Figure P2.71

Do this in the three different ways suggested by eq. (P2.71–1) and by the figure:

(i) By first convolving the two impulse responses and then convolving the result with $x(t)$.

(ii) By first convolving $x(t)$ with $u_1(t)$ and then convolving the result with $u(t)$.

(iii) By first convolving $x(t)$ with $u(t)$ and then convolving the result with $u_1(t)$.

(b) Repeat part (a) for

$$x(t) = e^{-t}$$

and

$$h(t) = e^{-t}u(t),$$
$$g(t) = u_1(t) + \delta(t).$$

(c) Do the same for

$$x[n] = \left(\frac{1}{2}\right)^n,$$

$$h[n] = \left(\frac{1}{2}\right)^n u[n],$$

$$g[n] = \delta[n] - \frac{1}{2}\delta[n-1].$$

Thus, in general, the associativity property of convolution holds if and only if the three expressions in eq. (P2.71–1) make sense (i.e., if and only if their interpretations in terms of LTI systems are meaningful). For example, in part (a) differentiating a constant and then integrating makes sense, but the process of integrating the constant from $t = -\infty$ and *then* differentiating does not, and it is only in such cases that associativity breaks down.

Closely related to the foregoing discussion is an issue involving inverse systems. Consider the LTI system with impulse response $h(t) = u(t)$. As we saw in part (a), there are inputs—specifically, $x(t) =$ nonzero constant—for which the output of this system is infinite, and thus, it is meaningless to consider the question of inverting such outputs to recover the input. However, if we limit ourselves to inputs that do yield finite outputs, that is, inputs which satisfy

$$\left|\int_{-\infty}^{t} x(\tau)\,d\tau\right| < \infty, \tag{P2.71–2}$$

then the system *is* invertible, and the LTI system with impulse response $u_1(t)$ is its inverse.

(d) Show that the LTI system with impulse response $u_1(t)$ is *not* invertible. (*Hint:* Find two different inputs that both yield zero output for all time.) However, show that the system is invertible if we limit ourselves to inputs that satisfy eq. (P2.71–2). [*Hint:* In Problem 1.44, we showed that an LTI system is invertible if no input other than $x(t) = 0$ yields an output that is zero for all time; are there two inputs $x(t)$ that satisfy eq. (P2.71–2) and that yield identically zero responses when convolved with $u_1(t)$?]

What we have illustrated in this problem is the following:

(1) If $x(t)$, $h(t)$, and $g(t)$ are three signals, and if $x(t) * g(t)$, $x(t) * h(t)$, and $h(t) * g(t)$ are *all* well defined and finite, then the associativity property, eq. (P2.71–1), holds.

(2) Let $h(t)$ be the impulse response of an LTI system, and suppose that the impulse response $g(t)$ of a second system has the property

$$h(t) * g(t) = \delta(t). \qquad \text{(P2.71–3)}$$

Then, from (1), *for* all inputs $x(t)$ for which $x(t) * h(t)$ and $x(t) * g(t)$ are both well defined and finite, the two cascades of systems depicted in Figure P2.71 act as the identity system, and thus, the two LTI systems can be regarded as inverses of one another. For example, if $h(t) = u(t)$ and $g(t) = u_1(t)$, then, as long as we restrict ourselves to inputs satisfying eq. (P2.71–2), we can regard these two systems as inverses.

Therefore, we see that the associativity property of eq. (P2.71–1) and the definition of LTI inverses as given in eq. (P2.71–3) are valid, as long as all convolutions that are involved are finite. As this is certainly the case in any realistic problem, we will in general use these properties without comment or qualification. Note that, although we have phrased most of our discussion in terms of continuous-time signals and systems, the same points can also be made in discrete time [as should be evident from part (c)].

2.72. Let $\delta_\Delta(t)$ denote the rectangular pulse of height $\frac{1}{\Delta}$ for $0 < t \leq \Delta$. Verify that

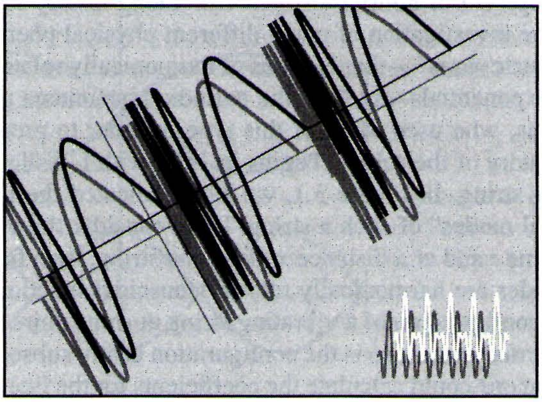
$$\frac{d}{dt}\delta_\Delta(t) = \frac{1}{\Delta}[\delta(t) - \delta(t - \Delta)].$$

2.73. Show by induction that

$$u_{-k}(t) = \frac{t^{k-1}}{(k-1)!}u(t) \text{ for } k = 1, 2, 3 \ldots$$

3

FOURIER SERIES REPRESENTATION OF PERIODIC SIGNALS



3.0 INTRODUCTION

The representation and analysis of LTI systems through the convolution sum as developed in Chapter 2 is based on representing signals as linear combinations of shifted impulses. In this and the following two chapters, we explore an alternative representation for signals and LTI systems. As in Chapter 2, the starting point for our discussion is the development of a representation of signals as linear combinations of a set of basic signals. For this alternative representation we use complex exponentials. The resulting representations are known as the continuous-time and discrete-time Fourier series and transform. As we will see, these can be used to construct broad and useful classes of signals.

We then proceed as we did in Chapter 2. That is, because of the superposition property, the response of an LTI system to any input consisting of a linear combination of basic signals is the same linear combination of the individual responses to each of the basic signals. In Chapter 2, these responses were all shifted versions of the unit impulse response, leading to the convolution sum or integral. As we will find in the current chapter, the response of an LTI system to a complex exponential also has a particularly simple form, which then provides us with another convenient representation for LTI systems and with another way in which to analyze these systems and gain insight into their properties.

In this chapter, we focus on the representation of continuous-time and discrete-time periodic signals referred to as the Fourier series. In Chapters 4 and 5, we extend the analysis to the Fourier transform representation of broad classes of aperiodic, finite energy signals. Together, these representations provide one of the most powerful and important sets of tools and insights for analyzing, designing, and understanding signals and LTI systems, and we devote considerable attention in this and subsequent chapters to exploring the uses of Fourier methods.

We begin in the next section with a brief historical perspective in order to provide some insight into the concepts and issues that we develop in more detail in the sections and chapters that follow.

3.1 A HISTORICAL PERSPECTIVE

The development of Fourier analysis has a long history involving a great many individuals and the investigation of many different physical phenomena.[1] The concept of using "trigonometric sums"—that is, sums of harmonically related sines and cosines or periodic complex exponentials—to describe periodic phenomena goes back at least as far as the Babylonians, who used ideas of this type in order to predict astronomical events.[2] The modern history of the subject begins in 1748 with L. Euler, who examined the motion of a vibrating string. In Figure 3.1, we have indicated the first few of what are known as the "normal modes" of such a string. If we consider the vertical deflection $f(t, x)$ of the string at time t and at a distance x along the string, then for any fixed instant of time, the normal modes are harmonically related sinusoidal functions of x. What Euler noted was that if the configuration of a vibrating string at some point in time is a linear combination of these normal modes, so is the configuration at any subsequent time. Furthermore, Euler showed that one could calculate the coefficients for the linear combination at the later time in a very straightforward manner from the coefficients at the earlier time. In doing this, Euler performed the same type of calculation as we will in the next section in deriving one of the properties of trigonometric sums that make them so useful for the analysis of LTI systems. Specifically, we will see that if the input to an LTI system is expressed as a linear combination of periodic complex exponentials or sinusoids, the output can also be expressed in this form, with coefficients that are related in a straightforward way to those of the input.

The property described in the preceding paragraph would not be particularly useful, unless it were true that a large class of interesting functions could be represented by linear combinations of complex exponentials. In the middle of the 18th century, this point was the subject of heated debate. In 1753, D. Bernoulli argued on physical grounds that all physical motions of a string could be represented by linear combinations of normal modes, but he did not pursue this mathematically, and his ideas were not widely accepted. In fact, Euler himself discarded trigonometric series, and in 1759 J. L. Lagrange strongly criticized the use of trigonometric series in the examination of vibrating strings. His criticism was based on his own belief that it was impossible to represent signals with corners (i.e., with discontinuous slopes) using trigonometric series. Since such a configuration arises from

---

[1] The historical material in this chapter was taken from the following references: I. Grattan-Guiness, *Joseph Fourier, 1768–1830* (Cambridge, MA: The MIT Press, 1972); G. F. Simmons, *Differential Equations: With Applications and Historical Notes* (New York: McGraw-Hill Book Company, 1972); C. Lanczos, *Discourse on Fourier Series* (London: Oliver and Boyd, 1966); R. E. Edwards, *Fourier Series: A Modern Introduction* (New York: Springer-Verlag, 2nd ed., 1970); and A. D. Aleksandrov, A. N. Kolmogorov, and M. A. Lavrent'ev, *Mathematics: Its Content, Methods, and Meaning,* trans. S. H. Gould, Vol. II; trans. K. Hirsch, Vol. III (Cambridge, MA: The MIT Press, 1969). Of these, Grattan-Guiness' work offers the most complete account of Fourier's life and contributions. Other references are cited in several places in the chapter.

[2] H. Dym and H. P. McKean, *Fourier Series and Integrals* (New York: Academic Press, 1972). This text and the book of Simmons cited in footnote 1 also contain discussions of the vibrating-string problem and its role in the development of Fourier analysis.

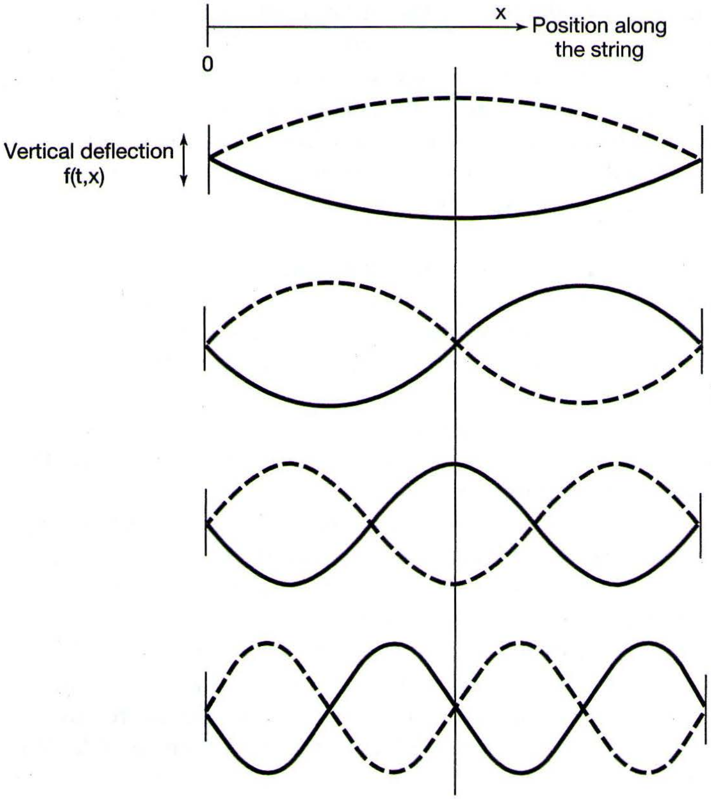


Figure 3.1 Normal modes of a vibrating string. (Solid lines indicate the configuration of each of these modes at some fixed instant of time, *t*.)

the plucking of a string (i.e., pulling it taut and then releasing it), Lagrange argued that trigonometric series were of very limited use.

It was in this somewhat hostile and skeptical environment that Jean Baptiste Joseph Fourier (Figure 3.2) presented his ideas half a century later. Fourier was born on March

Figure 3.2 Jean Baptiste Joseph Fourier [picture from J. B. J. Fourier, *Oeuvres de Fourier,* Vol. II (Paris: Gauthier-Villars et Fils, 1980)].

21, 1768, in Auxerre, France, and by the time of his entrance into the controversy concerning trigonometric series, he had already had a lifetime of experiences. His many contributions—in particular, those concerned with the series and transform that carry his name—are made even more impressive by the circumstances under which he worked. His revolutionary discoveries, although not completely appreciated during his own lifetime, have had a major impact on the development of mathematics and have been and still are of great importance in an extremely wide range of scientific and engineering disciplines.

In addition to his studies in mathematics, Fourier led an active political life. In fact, during the years that followed the French Revolution, his activities almost led to his downfall, as he narrowly avoided the guillotine on two separate occasions. Subsequently, Fourier became an associate of Napoleon Bonaparte, accompanied him on his expeditions to Egypt (during which time Fourier collected the information he would use later as the basis for his treatises on Egyptology), and in 1802 was appointed by Bonaparte to the position of prefect of a region of France centered in Grenoble. It was there, while serving as prefect, that Fourier developed his ideas on trigonometric series.

The physical motivation for Fourier's work was the phenomenon of heat propagation and diffusion. This in itself was a significant step in that most previous research in mathematical physics had dealt with rational and celestial mechanics. By 1807, Fourier had completed a work, Fourier had found series of harmonically related sinusoids to be useful in representing the temperature distribution through a body. In addition, he claimed that "any" periodic signal could be represented by such a series. While his treatment of this topic was significant, many of the basic ideas behind it had been discovered by others. Also, Fourier's mathematical arguments were still imprecise, and it remained for P. L. Dirichlet in 1829 to provide precise conditions under which a periodic signal could be represented by a Fourier series.[3] Thus, Fourier did not actually contribute to the mathematical theory of Fourier series. However, he did have the clear insight to see the potential for this series representation, and it was to a great extent his work and his claims that spurred much of the subsequent work on Fourier series. In addition, Fourier took this type of representation one very large step farther than any of his predecessors: He obtained a representation for *aperiodic* signals—not as weighted *sums* of harmonically related sinusoids—but as weighted *integrals* of sinusoids that are *not* all harmonically related. It is this extension from Fourier series to the Fourier integral or transform that is the focus of Chapters 4 and 5. Like the Fourier series, the Fourier transform remains one of the most powerful tools for the analysis of LTI systems.

Four distinguished mathematicians and scientists were appointed to examine the 1807 paper of Fourier. Three of the four—S. F. Lacroix, G. Monge, and P. S. de Laplace—were in favor of publication of the paper, but the fourth, J. L. Lagrange, remained adamant in rejecting trigonometric series, as he had done 50 years earlier. Because of Lagrange's vehement objections, Fourier's paper never appeared. After several other attempts to have his work accepted and published by the Institut de France, Fourier undertook the writing of another version of his work, which appeared as the text *Théorie analytique de la chaleur*.[4]

---

[3] Both S. D. Poisson and A. L. Cauchy had obtained results about the convergence of Fourier series before 1829, but Dirichlet's work represented such a significant extension of their results that he is usually credited with being the first to consider Fourier series convergence in a rigorous fashion.

[4] See J. B. J. Fourier, *The Analytical Theory of Heat*, trans. A. Freeman (New York: Dover, 1955).

This book was published in 1822, 15 years after Fourier had first presented his results to the Institut.

Toward the end of his life Fourier received some of the recognition he deserved, but the most significant tribute to him has been the enormous impact of his work on so many disciplines within the fields of mathematics, science, and engineering. The theory of integration, point-set topology, and eigenfunction expansions are just a few examples of topics in mathematics that have their roots in the analysis of Fourier series and integrals.[5] Furthermore, in addition to the original studies of vibration and heat diffusion, there are numerous other problems in science and engineering in which sinusoidal signals, and therefore Fourier series and transforms, play an important role. For example, sinusoidal signals arise naturally in describing the motion of the planets and the periodic behavior of the earth's climate. Alternating-current sources generate sinusoidal voltages and currents, and, as we will see, the tools of Fourier analysis enable us to analyze the response of an LTI system, such as a circuit, to such sinusoidal inputs. Also, as illustrated in Figure 3.3, waves in the ocean consist of the linear combination of sinusoidal waves with different spatial periods or *wavelengths*. Signals transmitted by radio and television stations are sinusoidal in nature as well, and as a quick perusal of any text on Fourier analysis will show, the range of applications in which sinusoidal signals arise and in which the tools of Fourier analysis are useful extends far beyond these few examples.

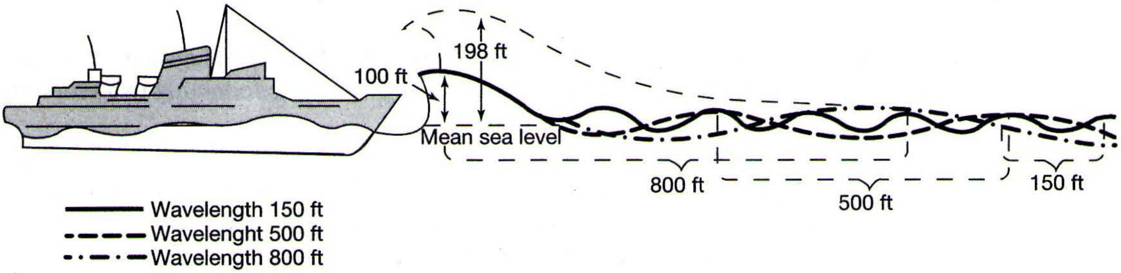


Figure 3.3 Ship encountering the superposition of three wave trains, each with a different spatial period. When these waves reinforce one another, a very large wave can result. In more severe seas, a giant wave indicated by the dotted line could result. Whether such a reinforcement occurs at any location depends upon the relative phases of the components that are superposed. [Adapted from an illustration by P. Mion in "Nightmare Waves Are All Too Real to Deepwater Sailors," by P. Britton, *Smithsonian* 8 (February 1978), pp. 64–65].

While many of the applications in the preceding paragraph, as well as the original work of Fourier and his contemporaries on problems of mathematical physics, focus on phenomena in continuous time, the tools of Fourier analysis for discrete-time signals and systems have their own distinct historical roots and equally rich set of applications. In particular, discrete-time concepts and methods are fundamental to the discipline of numerical analysis. Formulas for the processing of discrete sets of data points to produce numerical approximations for interpolation, integration, and differentiation were being investigated as early as the time of Newton in the 1600s. In addition, the problem of predicting the motion of a heavenly body, given a sequence of observations of the body, spurred the investigation of harmonic time series in the 18th and 19th centuries by eminent scientists and mathematicians, including Gauss, and thus provided a second setting in which much of the initial work was done on discrete-time signals and systems.

---

[5] For more on the impact of Fourier's work on mathematics, see W. A. Coppel, "J. B. Fourier—on the occasion of His Two Hundredth Birthday," *American Mathematical Monthly*, 76 (1969), 468–83.

In the mid-1960s an algorithm, now known as the fast Fourier transform, or FFT, was introduced. This algorithm, which was independently discovered by Cooley and Tukey in 1965, also has a considerable history and can, in fact, be found in Gauss' notebooks.[6] What made its modern discovery so important was the fact that the FFT proved to be perfectly suited for efficient digital implementation, and it reduced the time required to compute transforms by orders of magnitude. With this tool, many interesting but previously impractical ideas utilizing the discrete-time Fourier series and transform suddenly became practical, and the development of discrete-time signal and system analysis techniques moved forward at an accelerated pace.

What has emerged out of this long history is a powerful and cohesive framework for the analysis of continuous-time and discrete-time signals and systems and an extraordinarily broad array of existing and potential applications. In this and the following chapters, we will develop the basic tools of that framework and examine some of its important implications.

3.2 THE RESPONSE OF LTI SYSTEMS TO COMPLEX EXPONENTIALS

As we indicated in Section 3.0, it is advantageous in the study of LTI systems to represent signals as linear combinations of basic signals that possess the following two properties:

1. The set of basic signals can be used to construct a broad and useful class of signals.
2. The response of an LTI system to each signal should be simple enough in structure to provide us with a convenient representation for the response of the system to any signal constructed as a linear combination of the basic signals.

Much of the importance of Fourier analysis results from the fact that both of these properties are provided by the set of complex exponential signals in continuous and discrete time—i.e., signals of the form e^{st} in continuous time and z^n in discrete time, where s and z are complex numbers. In subsequent sections of this and the following two chapters, we will examine the first property in some detail. In this section, we focus on the second property and, in this way, provide motivation for the use of Fourier series and transforms in the analysis of LTI systems.

The importance of complex exponentials in the study of LTI systems stems from the fact that the response of an LTI system to a complex exponential input is the same complex exponential with only a change in amplitude; that is,

$$\text{continuous time: } e^{st} \longrightarrow H(s)e^{st}, \tag{3.1}$$

$$\text{discrete time: } z^n \longrightarrow H(z)z^n, \tag{3.2}$$

where the complex amplitude factor $H(s)$ or $H(z)$ will in general be a function of the complex variable s or z. A signal for which the system output is a (possibly complex) constant times the input is referred to as an *eigenfunction* of the system, and the amplitude factor is referred to as the system's *eigenvalue*.

---

[6] M. T. Heideman, D. H. Johnson, and C. S. Burrus, "Gauss and the History of the Fast Fourier Transform," *The IEEE ASSP Magazine I* (1984), pp. 14–21.

To show that complex exponentials are indeed eigenfunctions of LTI systems, let us consider a continuous-time LTI system with impulse response $h(t)$. For an input $x(t)$, we can determine the output through the use of the convolution integral, so that with $x(t) = e^{st}$

$$\begin{aligned} y(t) &= \int_{-\infty}^{+\infty} h(\tau)x(t-\tau)\,d\tau \\ &= \int_{-\infty}^{+\infty} h(\tau)e^{s(t-\tau)}\,d\tau. \end{aligned} \tag{3.3}$$

Expressing $e^{s(t-\tau)}$ as $e^{st}e^{-s\tau}$, and noting that e^{st} can be moved outside the integral, we see that eq. (3.3) becomes

$$y(t) = e^{st}\int_{-\infty}^{+\infty} h(\tau)e^{-s\tau}\,d\tau. \tag{3.4}$$

Assuming that the integral on the right-hand side of eq. (3.4) converges, the response to e^{st} is of the form

$$y(t) = H(s)e^{st}, \tag{3.5}$$

where $H(s)$ is a complex constant whose value depends on s and which is related to the system impulse response by

$$H(s) = \int_{-\infty}^{+\infty} h(\tau)e^{-s\tau}\,d\tau. \tag{3.6}$$

Hence, we have shown that complex exponentials are eigenfunctions of LTI systems. The constant $H(s)$ for a specific value of s is then the eigenvalue associated with the eigenfunction e^{st}.

In an exactly parallel manner, we can show that complex exponential sequences are eigenfunctions of discrete-time LTI systems. That is, suppose that an LTI system with impulse response $h[n]$ has as its input the sequence

$$x[n] = z^n, \tag{3.7}$$

where z is a complex number. Then the output of the system can be determined from the convolution sum as

$$\begin{aligned} y[n] &= \sum_{k=-\infty}^{+\infty} h[k]x[n-k] \\ &= \sum_{k=-\infty}^{+\infty} h[k]z^{n-k} = z^n \sum_{k=-\infty}^{+\infty} h[k]z^{-k}. \end{aligned} \tag{3.8}$$

From this expression, we see that if the input $x[n]$ is the complex exponential given by eq. (3.7), then, assuming that the summation on the right-hand side of eq. (3.8) converges, the output is the same complex exponential multiplied by a constant that depends on the value of z. That is,

$$y[n] = H(z)z^n, \tag{3.9}$$

where

$$H(z) = \sum_{k=-\infty}^{+\infty} h[k]z^{-k}. \tag{3.10}$$

Consequently, as in the continuous-time case, complex exponentials are eigenfunctions of discrete-time LTI systems. The constant $H(z)$ for a specified value of z is the eigenvalue associated with the eigenfunction z^n.

For the analysis of LTI systems, the usefulness of decomposing more general signals in terms of eigenfunctions can be seen from an example. Let $x(t)$ correspond to a linear combination of three complex exponentials; that is,

$$x(t) = a_1 e^{s_1 t} + a_2 e^{s_2 t} + a_3 e^{s_3 t}. \tag{3.11}$$

From the eigenfunction property, the response to each separately is

$$\begin{aligned} a_1 e^{s_1 t} &\longrightarrow a_1 H(s_1) e^{s_1 t}, \\ a_2 e^{s_2 t} &\longrightarrow a_2 H(s_2) e^{s_2 t}, \\ a_3 e^{s_3 t} &\longrightarrow a_3 H(s_3) e^{s_3 t}, \end{aligned}$$

and from the superposition property the response to the sum is the sum of the responses, so that

$$y(t) = a_1 H(s_1) e^{s_1 t} + a_2 H(s_2) e^{s_2 t} + a_3 H(s_3) e^{s_3 t}. \tag{3.12}$$

More generally, in continuous time, eq. (3.5), together with the superposition property, implies that the representation of signals as a linear combination of complex exponentials leads to a convenient expression for the response of an LTI system. Specifically, if the input to a continuous-time LTI system is represented as a linear combination of complex exponentials, that is, if

$$x(t) = \sum_k a_k e^{s_k t}, \tag{3.13}$$

then the output will be

$$y(t) = \sum_k a_k H(s_k) e^{s_k t}. \tag{3.14}$$

In an exactly analogous manner, if the input to a discrete-time LTI system is represented as a linear combination of complex exponentials, that is, if

$$x[n] = \sum_k a_k z_k^n, \tag{3.15}$$

then the output will be

$$y[n] = \sum_k a_k H(z_k) z_k^n. \tag{3.16}$$

In other words, for both continuous time and discrete time, if the input to an LTI system is represented as a linear combination of complex exponentials, then the output can also be represented as a linear combination of the same complex exponential signals. Each coefficient in this representation of the output is obtained as the product of the corresponding coefficient a_k of the input and the system's eigenvalue $H(s_k)$ or $H(z_k)$ associated with the eigenfunction $e^{s_k t}$ or z_k^n, respectively. It was precisely this fact that Euler discovered for the problem of the vibrating string, that Gauss and others used in the analysis of time series, and that motivated Fourier and others after him to consider the question of how broad a class of signals could be represented as a linear combination of complex exponentials. In the next few sections we examine this question for periodic signals, first in continuous time and then in discrete time, and in Chapters 4 and 5 we consider the extension of these representations to aperiodic signals. Although in general, the variables s and z in eqs. (3.1)–(3.16) may be arbitrary complex numbers, Fourier analysis involves restricting our attention to particular forms for these variables. In particular, in continuous time we focus on purely imaginary values of s—i.e., $s = j\omega$—and thus, we consider only complex exponentials of the form $e^{j\omega t}$. Similarly, in discrete time we restrict the range of values of z to those of unit magnitude—i.e., $z = e^{j\omega}$—so that we focus on complex exponentials of the form $e^{j\omega n}$.

Example 3.1

As an illustration of eqs. (3.5) and (3.6), consider an LTI system for which the input $x(t)$ and output $y(t)$ are related by a time shift of 3, i.e.,

$$y(t) = x(t-3). \tag{3.17}$$

If the input to this system is the complex exponential signal $x(t) = e^{j2t}$, then, from eq. (3.17),

$$y(t) = e^{j2(t-3)} = e^{-j6}e^{j2t}. \tag{3.18}$$

Equation (3.18) is in the form of eq. (3.5), as we would expect, since e^{j2t} is an eigenfunction. The associated eigenvalue is $H(j2) = e^{-j6}$. It is straightforward to confirm eq. (3.6) for this example. Specifically, from eq. (3.17), the impulse response of the system is $h(t) = \delta(t-3)$. Substituting into eq. (3.6), we obtain

$$H(s) = \int_{-\infty}^{+\infty} \delta(\tau - 3)e^{-s\tau}\,d\tau = e^{-3s},$$

so that $H(j2) = e^{-j6}$.

As a second example, in this case illustrating eqs. (3.11) and (3.12), consider the input signal $x(t) = \cos(4t) + \cos(7t)$. From eq. (3.17), $y(t)$ will of course be

$$y(t) = \cos(4(t-3)) + \cos(7(t-3)). \tag{3.19}$$

To see that this will also result from eq. (3.12), we first expand $x(t)$ using Euler's relation:

$$x(t) = \tfrac{1}{2}e^{j4t} + \tfrac{1}{2}e^{-j4t} + \tfrac{1}{2}e^{j7t} + \tfrac{1}{2}e^{-j7t}. \tag{3.20}$$

From eqs. (3.11) and (3.12),

$$y(t) = \tfrac{1}{2}e^{-j12}e^{j4t} + \tfrac{1}{2}e^{j12}e^{-j4t} + \tfrac{1}{2}e^{-j21}e^{j7t} + \tfrac{1}{2}e^{j21}e^{-j7t},$$

or

$$\begin{aligned} y(t) &= \tfrac{1}{2}e^{j4(t-3)} + \tfrac{1}{2}e^{-j4(t-3)} + \tfrac{1}{2}e^{j7(t-3)} + \tfrac{1}{2}e^{-j7(t-3)} \\ &= \cos(4(t-3)) + \cos(7(t-3)). \end{aligned}$$

For this simple example, multiplication of each periodic exponential component of $x(t)$—for example, $\frac{1}{2}e^{j4t}$—by the corresponding eigenvalue—e.g., $H(j4) = e^{-j12}$—effectively causes the input component to shift in time by 3. Obviously, in this case we can determine $y(t)$ in eq. (3.19) by inspection rather than by employing eqs. (3.11) and (3.12). However, as we will see, the general property embodied in eqs. (3.11) and (3.12) not only allows us to calculate the responses of more complex LTI systems, but also provides the basis for the frequency domain representation and analysis of LTI systems.

3.3 FOURIER SERIES REPRESENTATION OF CONTINUOUS-TIME PERIODIC SIGNALS

3.3.1 Linear Combinations of Harmonically Related Complex Exponentials

As defined in Chapter 1, a signal is periodic if, for some positive value of T,

$$x(t) = x(t+T) \quad \text{for all } t. \tag{3.21}$$

The fundamental period of $x(t)$ is the minimum positive, nonzero value of T for which eq. (3.21) is satisfied, and the value $\omega_0 = 2\pi/T$ is referred to as the fundamental frequency.

In Chapter 1 we also introduced two basic periodic signals, the sinusoidal signal

$$x(t) = \cos\omega_0 t \tag{3.22}$$

and the periodic complex exponential

$$x(t) = e^{j\omega_0 t}. \tag{3.23}$$

Both of these signals are periodic with fundamental frequency ω_0 and fundamental period $T = 2\pi/\omega_0$. Associated with the signal in eq. (3.23) is the set of *harmonically related* complex exponentials

$$\phi_k(t) = e^{jk\omega_0 t} = e^{jk(2\pi/T)t}, \quad k = 0, \pm1, \pm2, \ldots. \tag{3.24}$$

Each of these signals has a fundamental frequency that is a multiple of ω_0, and therefore, each is periodic with period T (although for $|k| \geq 2$, the fundamental period of $\phi_k(t)$ is a fraction of T). Thus, a linear combination of harmonically related complex exponentials of the form

$$x(t) = \sum_{k=-\infty}^{+\infty} a_k e^{jk\omega_0 t} = \sum_{k=-\infty}^{+\infty} a_k e^{jk(2\pi/T)t} \tag{3.25}$$

is also periodic with period T. In eq. (3.25), the term for $k = 0$ is a constant. The terms for $k = +1$ and $k = -1$ both have fundamental frequency equal to ω_0 and are collectively referred to as the *fundamental components* or the *first harmonic components.* The two terms for $k = +2$ and $k = -2$ are periodic with half the period (or, equivalently, twice the frequency) of the fundamental components and are referred to as the *second harmonic components.* More generally, the components for $k = +N$ and $k = -N$ are referred to as the Nth harmonic components.

The representation of a periodic signal in the form of eq. (3.25) is referred to as the *Fourier series* representation. Before developing the properties of this representation, let us consider an example.

Example 3.2

Consider a periodic signal $x(t)$, with fundamental frequency 2π, that is expressed in the form of eq. (3.25) as

$$x(t) = \sum_{k=-3}^{+3} a_k e^{jk2\pi t}, \tag{3.26}$$

where

$$\begin{aligned} a_0 &= 1, \\ a_1 &= a_{-1} = \frac{1}{4}, \\ a_2 &= a_{-2} = \frac{1}{2}, \\ a_3 &= a_{-3} = \frac{1}{3}. \end{aligned}$$

Rewriting eq. (3.26) and collecting each of the harmonic components which have the same fundamental frequency, we obtain

$$\begin{aligned} x(t) = 1 &+ \frac{1}{4}(e^{j2\pi t} + e^{-j2\pi t}) + \frac{1}{2}(e^{j4\pi t} + e^{-j4\pi t}) \\ &+ \frac{1}{3}(e^{j6\pi t} + e^{-j6\pi t}). \end{aligned} \tag{3.27}$$

Equivalently, using Euler's relation, we can write $x(t)$ in the form

$$x(t) = 1 + \frac{1}{2}\cos 2\pi t + \cos 4\pi t + \frac{2}{3}\cos 6\pi t. \tag{3.28}$$

In Figure 3.4, we illustrate graphically how the signal $x(t)$ is built up from its harmonic components.

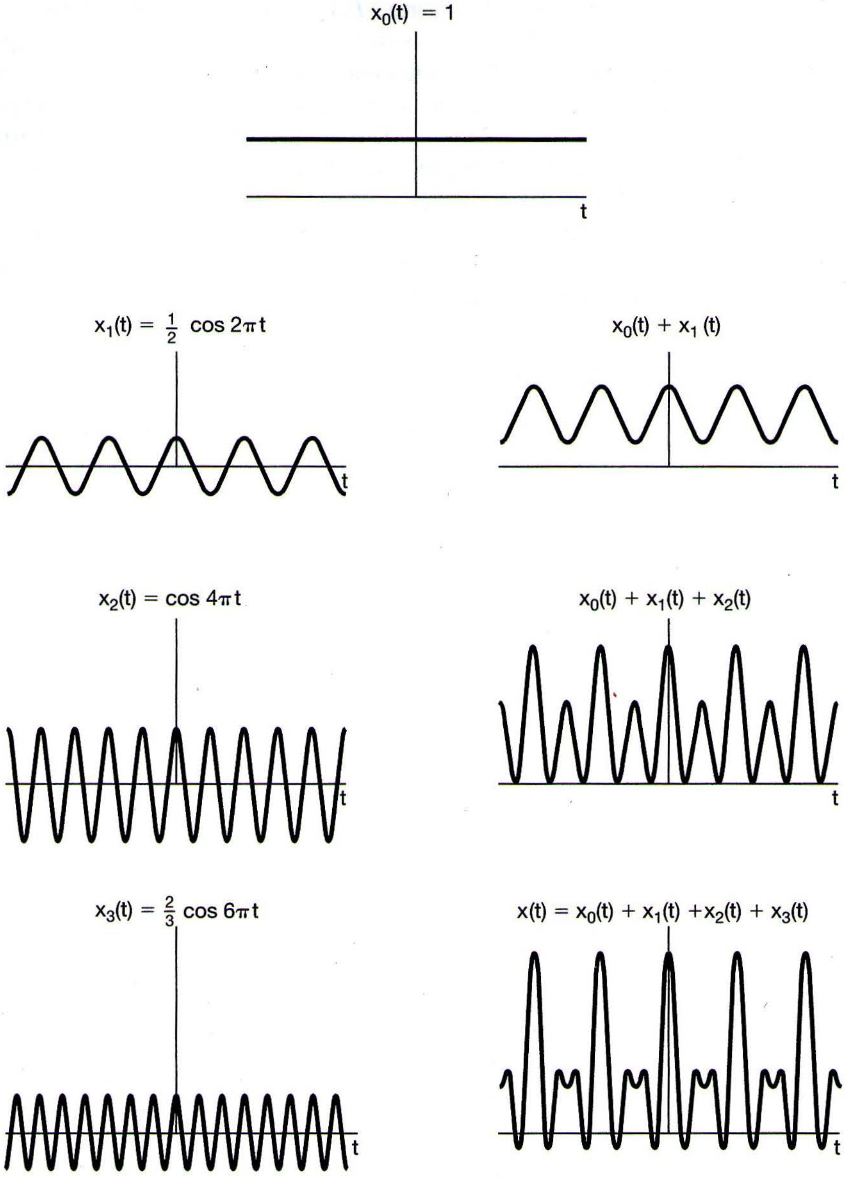


Figure 3.4 Construction of the signal $x(t)$ in Example 3.2 as a linear combination of harmonically related sinusoidal signals.

Equation (3.28) is an example of an alternative form for the Fourier series of real periodic signals. Specifically, suppose that $x(t)$ is real and can be represented in the form of eq. (3.25). Then, since $x^*(t) = x(t)$, we obtain

$$x(t) = \sum_{k=-\infty}^{+\infty} a_k^* e^{-jk\omega_0 t}.$$

Replacing k by $-k$ in the summation, we have

$$x(t) = \sum_{k=-\infty}^{+\infty} a_{-k}^* e^{jk\omega_0 t},$$

which, by comparison with eq. (3.25), requires that $a_k = a_{-k}^*$, or equivalently, that

$$a_k^* = a_{-k}. \tag{3.29}$$

Note that this is the case in Example 3.2, where the a_k's are in fact real and $a_k = a_{-k}$.

To derive the alternative forms of the Fourier series, we first rearrange the summation in eq. (3.25) as

$$x(t) = a_0 + \sum_{k=1}^{\infty} [a_k e^{jk\omega_0 t} + a_{-k} e^{-jk\omega_0 t}].$$

Substituting a_k^* for a_{-k} from eq. (3.29), we obtain

$$x(t) = a_0 + \sum_{k=1}^{\infty} [a_k e^{jk\omega_0 t} + a_k^* e^{-jk\omega_0 t}].$$

Since the two terms inside the summation are complex conjugates of each other, this can be expressed as

$$x(t) = a_0 + \sum_{k=1}^{\infty} 2\mathcal{R}e\{a_k e^{jk\omega_0 t}\}. \tag{3.30}$$

If a_k is expressed in polar form as

$$a_k = A_k e^{j\theta_k},$$

then eq. (3.30) becomes

$$x(t) = a_0 + \sum_{k=1}^{\infty} 2\mathcal{R}e\{A_k e^{j(k\omega_0 t + \theta_k)}\}.$$

That is,

$$x(t) = a_0 + 2\sum_{k=1}^{\infty} A_k \cos(k\omega_0 t + \theta_k). \tag{3.31}$$

Equation (3.31) is one commonly encountered form for the Fourier series of real periodic signals in continuous time. Another form is obtained by writing a_k in rectangular form as

$$a_k = B_k + jC_k,$$

where B_k and C_k are both real. With this expression for a_k, eq. (3.30) takes the form

$$x(t) = a_0 + 2\sum_{k=1}^{\infty} [B_k \cos k\omega_0 t - C_k \sin k\omega_0 t]. \tag{3.32}$$

In Example 3.2 the a_k's are all real, so that $a_k = A_k = B_k$, and therefore, both representations, eqs. (3.31) and (3.32), reduce to the same form, eq. (3.28).

Thus, for real periodic functions, the Fourier series in terms of complex exponentials, as given in eq. (3.25), is mathematically equivalent to either of the two forms in eqs. (3.31) and (3.32) that use trigonometric functions. Although the latter two are common forms for Fourier series,[7] the complex exponential form of eq. (3.25) is particularly convenient for our purposes, so we will use that form almost exclusively.

Equation (3.29) illustrates one of many properties associated with Fourier series. These properties are often quite useful in gaining insight and for computational purposes, and in Section 3.5 we collect together the most important of them. The derivation of several of them is considered in problems at the end of the chapter. In Section 4.3, we also will develop the majority of the properties within the broader context of the Fourier transform.

3.3.2 Determination of the Fourier Series Representation of a Continuous-time Periodic Signal

Assuming that a given periodic signal can be represented with the series of eq. (3.25), we need a procedure for determining the coefficients a_k. Multiplying both sides of eq. (3.25) by $e^{-jn\omega_0 t}$, we obtain

$$x(t)e^{-jn\omega_0 t} = \sum_{k=-\infty}^{+\infty} a_k e^{jk\omega_0 t} e^{-jn\omega_0 t}. \tag{3.33}$$

Integrating both sides from 0 to $T = 2\pi/\omega_0$, we have

$$\int_0^T x(t)e^{-jn\omega_0 t}\,dt = \int_0^T \sum_{k=-\infty}^{+\infty} a_k e^{jk\omega_0 t} e^{-jn\omega_0 t}\,dt.$$

Here, T is the fundamental period of $x(t)$, and consequently, we are integrating over one period. Interchanging the order of integration and summation yields

$$\int_0^T x(t)e^{-jn\omega_0 t}\,dt = \sum_{k=-\infty}^{+\infty} a_k \left[\int_0^T e^{j(k-n)\omega_0 t}\,dt\right]. \tag{3.34}$$

The evaluation of the bracketed integral is straightforward. Rewriting this integral using Euler's formula, we obtain

$$\int_0^T e^{j(k-n)\omega_0 t}\,dt = \int_0^T \cos(k-n)\omega_0 t\,dt + j\int_0^T \sin(k-n)\omega_0 t\,dt. \tag{3.35}$$

For $k \neq n$, $\cos(k-n)\omega_0 t$ and $\sin(k-n)\omega_0 t$ are periodic sinusoids with fundamental period $(T/|k-n|)$. Therefore, in eq. (3.35), we are integrating over an interval (of length T) that is an integral number of periods of these signals. Since the integral may be viewed as measuring the total area under the functions over the interval, we see that for $k \neq n$, both of the integrals on the right-hand side of eq. (3.35) are zero. For $k = n$, the integrand on the left-hand side of eq. (3.35) equals 1, and thus, the integral equals T. In sum, we then have

$$\int_0^T e^{j(k-n)\omega_0 t}\,dt = \begin{cases} T, & k = n \\ 0, & k \neq n \end{cases},$$

---

[7] In fact, in his original work, Fourier used the sine-cosine form of the Fourier series given in eq. (3.32).

and consequently, the right-hand side of eq. (3.34) reduces to Ta_n. Therefore,

$$a_n = \frac{1}{T}\int_0^T x(t)e^{-jn\omega_0 t}\,dt, \tag{3.36}$$

which provides the equation for determining the coefficients. Furthermore, note that in evaluating eq. (3.35), the only fact that we used concerning the interval of integration was that we were integrating over an interval of length T, which is an integral number of periods of $\cos(k-n)\omega_0 t$ and $\sin(k-n)\omega_0 t$. Therefore, we will obtain the same result if we integrate over any interval of length T. That is, if we denote integration over *any* interval of length T by $\int_T$, we have

$$\int_T e^{j(k-n)\omega_0 t}\,dt = \begin{cases} T, & k = n \\ 0, & k \neq n \end{cases},$$

and consequently,

$$a_n = \frac{1}{T}\int_T x(t)e^{-jn\omega_0 t}\,dt. \tag{3.37}$$

To summarize, if $x(t)$ has a Fourier series representation [i.e., if it can be expressed as a linear combination of harmonically related complex exponentials in the form of eq. (3.25)], then the coefficients are given by eq. (3.37). This pair of equations, then, defines the Fourier series of a periodic continuous-time signal:

$$x(t) = \sum_{k=-\infty}^{+\infty} a_k e^{jk\omega_0 t} = \sum_{k=-\infty}^{+\infty} a_k e^{jk(2\pi/T)t}, \tag{3.38}$$

$$a_k = \frac{1}{T}\int_T x(t)e^{-jk\omega_0 t}\,dt = \frac{1}{T}\int_T x(t)e^{-jk(2\pi/T)t}\,dt. \tag{3.39}$$

Here, we have written equivalent expressions for the Fourier series in terms of the fundamental frequency ω_0 and the fundamental period T. Equation (3.38) is referred to as the *synthesis* equation and eq. (3.39) as the *analysis* equation. The set of coefficients $\{a_k\}$ are often called the *Fourier series coefficients* or the *spectral coefficients* of $x(t)$.[8] These complex coefficients measure the portion of the signal $x(t)$ that is at each harmonic of the fundamental component. The coefficient a_0 is the dc or constant component of $x(t)$ and is given by eq. (3.39) with $k = 0$. That is,

$$a_0 = \frac{1}{T}\int_T x(t)\,dt, \tag{3.40}$$

which is simply the average value of $x(t)$ over one period.

Equations (3.38) and (3.39) were known to both Euler and Lagrange in the middle of the 18th century. However, they discarded this line of analysis without having examined the question of how large a class of periodic signals could, in fact, be represented in such a fashion. Before we turn to this question in the next section, let us illustrate the continuous-time Fourier series by means of a few examples.

---

[8] The term "spectral coefficient" is derived from problems such as the spectroscopic decomposition of light into spectral lines (i.e., into its elementary components at different frequencies). The intensity of any line in such a decomposition is a direct measure of the fraction of the total light energy at the frequency corresponding to the line.

Example 3.3

Consider the signal

$$x(t) = \sin \omega_0 t,$$

whose fundamental frequency is ω_0. One approach to determining the Fourier series coefficients for this signal is to apply eq. (3.39). For this simple case, however, it is easier to expand the sinusoidal signal as a linear combination of complex exponentials and identify the Fourier series coefficients by inspection. Specifically, we can express $\sin \omega_0 t$ as

$$\sin \omega_0 t = \frac{1}{2j} e^{j\omega_0 t} - \frac{1}{2j} e^{-j\omega_0 t}.$$

Comparing the right-hand sides of this equation and eq. (3.38), we obtain

$$a_1 = \frac{1}{2j}, \qquad a_{-1} = -\frac{1}{2j},$$
$$a_k = 0, \qquad k \neq +1 \text{ or } -1.$$

Example 3.4

Let

$$x(t) = 1 + \sin \omega_0 t + 2 \cos \omega_0 t + \cos\left(2\omega_0 t + \frac{\pi}{4}\right),$$

which has fundamental frequency ω_0. As with Example 3.3, we can again expand $x(t)$ directly in terms of complex exponentials, so that

$$x(t) = 1 + \frac{1}{2j}[e^{j\omega_0 t} - e^{-j\omega_0 t}] + [e^{j\omega_0 t} + e^{-j\omega_0 t}] + \frac{1}{2}[e^{j(2\omega_0 t + \pi/4)} + e^{-j(2\omega_0 t + \pi/4)}].$$

Collecting terms, we obtain

$$x(t) = 1 + \left(1 + \frac{1}{2j}\right)e^{j\omega_0 t} + \left(1 - \frac{1}{2j}\right)e^{-j\omega_0 t} + \left(\frac{1}{2}e^{j(\pi/4)}\right)e^{j2\omega_0 t} + \left(\frac{1}{2}e^{-j(\pi/4)}\right)e^{-j2\omega_0 t}.$$

Thus, the Fourier series coefficients for this example are

$$a_0 = 1,$$
$$a_1 = \left(1 + \frac{1}{2j}\right) = 1 - \frac{1}{2}j,$$
$$a_{-1} = \left(1 - \frac{1}{2j}\right) = 1 + \frac{1}{2}j,$$
$$a_2 = \frac{1}{2}e^{j(\pi/4)} = \frac{\sqrt{2}}{4}(1 + j),$$

$$a_{-2} = \frac{1}{2}e^{-j(\pi/4)} = \frac{\sqrt{2}}{4}(1-j),$$
$$a_k = 0, \ |k| > 2.$$

In Figure 3.5, we show a bar graph of the magnitude and phase of a_k.

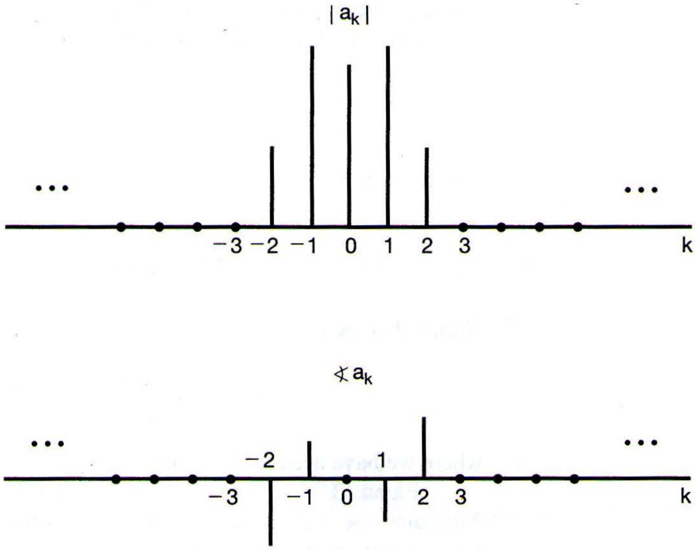


Figure 3.5 Plots of the magnitude and phase of the Fourier coefficients of the signal considered in Example 3.4.

Example 3.5

The periodic square wave, sketched in Figure 3.6 and defined over one period as

$$x(t) = \begin{cases} 1, & |t| < T_1 \\ 0, & T_1 < |t| < T/2 \end{cases}, \tag{3.41}$$

is a signal that we will encounter a number of times throughout this book. This signal is periodic with fundamental period T and fundamental frequency $\omega_0 = 2\pi/T$.

To determine the Fourier series coefficients for $x(t)$, we use eq. (3.39). Because of the symmetry of $x(t)$ about $t = 0$, it is convenient to choose $-T/2 \leq t < T/2$ as the

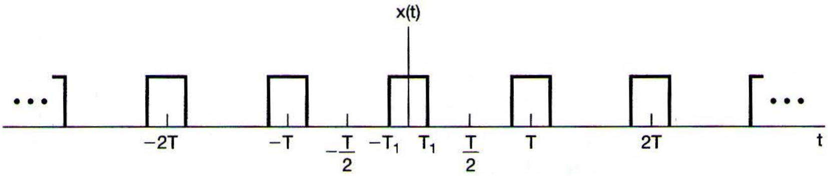


Figure 3.6 Periodic square wave.

interval over which the integration is performed, although any interval of length T is equally valid and thus will lead to the same result. Using these limits of integration and substituting from eq. (3.41), we have first, for $k = 0$,

$$a_0 = \frac{1}{T}\int_{-T_1}^{T_1} dt = \frac{2T_1}{T}. \tag{3.42}$$

As mentioned previously, a_0 is interpreted to be the average value of $x(t)$, which in this case equals the fraction of each period during which $x(t) = 1$. For $k \neq 0$, we obtain

$$a_k = \frac{1}{T}\int_{-T_1}^{T_1} e^{-jk\omega_0 t}\,dt = -\frac{1}{jk\omega_0 T}e^{-jk\omega_0 t}\Big|_{-T_1}^{T_1},$$

which we may rewrite as

$$a_k = \frac{2}{k\omega_0 T}\left[\frac{e^{jk\omega_0 T_1} - e^{-jk\omega_0 T_1}}{2j}\right]. \tag{3.43}$$

Noting that the term in brackets is $\sin k\omega_0 T_1$, we can express the coefficients a_k as

$$a_k = \frac{2\sin(k\omega_0 T_1)}{k\omega_0 T} = \frac{\sin(k\omega_0 T_1)}{k\pi}, \quad k \neq 0, \tag{3.44}$$

where we have used the fact that $\omega_0 T = 2\pi$.

Figure 3.7 is a bar graph of the Fourier series coefficients for this example. In particular, the coefficients are plotted for a fixed value of T_1 and several values of T. For this specific example, the Fourier coefficients are real, and consequently, they can be depicted graphically with only a single graph. More generally, of course, the Fourier coefficients are complex, so that two graphs, corresponding to the real and imaginary parts, or magnitude and phase, of each coefficient, would be required. For $T = 4T_1$, $x(t)$ is a square wave that is unity for half the period and zero for half the period. In this case, $\omega_0 T_1 = \pi/2$, and from eq. (3.44),

$$a_k = \frac{\sin(\pi k/2)}{k\pi}, \quad k \neq 0, \tag{3.45}$$

while

$$a_0 = \frac{1}{2}. \tag{3.46}$$

From eq. (3.45), $a_k = 0$ for k even and nonzero. Also, $\sin(\pi k/2)$ alternates between ± 1 for successive odd values of k. Therefore,

$$\begin{aligned} a_1 &= a_{-1} = \frac{1}{\pi}, \\ a_3 &= a_{-3} = -\frac{1}{3\pi}, \\ a_5 &= a_{-5} = \frac{1}{5\pi}, \\ &\ \ \vdots \end{aligned}$$

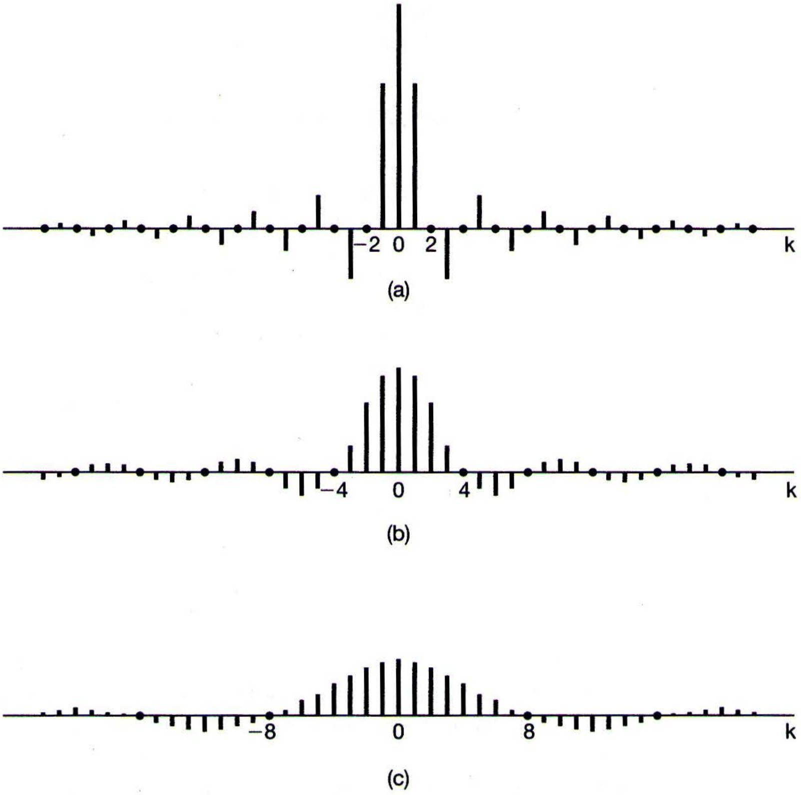


Figure 3.7 Plots of the scaled Fourier series coefficients Ta_k for the periodic square wave with T_1 fixed and for several values of T: (*a*) $T = 4T_1$; (*b*) $T = 8T_1$; (*c*) $T = 16T_1$. The coefficients are regularly spaced samples of the envelope $(2\sin\omega T_1)/\omega$, where the spacing between samples, $2\pi/T$, decreases as T increases.

3.4 CONVERGENCE OF THE FOURIER SERIES

Although Euler and Lagrange would have been happy with the results of Examples 3.3 and 3.4, they would have objected to Example 3.5, since $x(t)$ is discontinuous while each of its harmonic components is continuous. Fourier, on the other hand, considered the same example and maintained that the Fourier series representation of the square wave is valid. In fact, Fourier maintained that *any* periodic signal could be represented by a Fourier series. Although this is not quite true, it *is* true that Fourier series can be used to represent an extremely large class of periodic signals, including the square wave and all other periodic signals with which we will be concerned in this book and which are of interest in practice.

To gain an understanding of the square-wave example and, more generally, of the question of the validity of Fourier series representations, let us examine the problem of approximating a given periodic signal $x(t)$ by a linear combination of a finite number of harmonically related complex exponentials—that is, by a finite series of the form

$$x_N(t) = \sum_{k=-N}^{N} a_k e^{jk\omega_0 t}. \tag{3.47}$$

Let $e_N(t)$ denote the approximation error; that is,

$$e_N(t) = x(t) - x_N(t) = x(t) - \sum_{k=-N}^{+N} a_k e^{jk\omega_0 t}. \tag{3.48}$$

In order to determine how good any particular approximation is, we need to specify a quantitative measure of the size of the approximation error. The criterion that we will use is the energy in the error over one period:

$$E_N = \int_T |e_N(t)|^2 \, dt. \tag{3.49}$$

As shown in Problem 3.66, the particular choice for the coefficients in eq. (3.47) that minimize the energy in the error is

$$a_k = \frac{1}{T} \int_T x(t) e^{-jk\omega_0 t} \, dt. \tag{3.50}$$

Comparing eqs. (3.50) and (3.39), we see that eq. (3.50) is identical to the expression used to determine the Fourier series coefficients. Thus, if $x(t)$ has a Fourier series representation, the best approximation using only a finite number of harmonically related complex exponentials is obtained by truncating the Fourier series to the desired number of terms. As N increases, new terms are added and E_N decreases. If, in fact, $x(t)$ has a Fourier series representation, then the limit of E_N as $N \to \infty$ is zero.

Let us turn now to the question of when a periodic signal $x(t)$ does in fact have a Fourier series representation. Of course, for any signal, we can attempt to obtain a set of Fourier coefficients through the use of eq. (3.39). However, in some cases, the integral in eq. (3.39) may diverge; that is, the value obtained for some of the a_k may be infinite. Moreover, even if all of the coefficients obtained from eq. (3.39) are finite, when these coefficients are substituted into the synthesis equation (3.38), the resulting infinite series may not converge to the original signal $x(t)$.

Fortunately, there are no convergence difficulties for large classes of periodic signals. For example, every continuous periodic signal has a Fourier series representation for which the energy E_N in the approximation error approaches 0 as N goes to ∞. This is also true for many discontinuous signals. Since we will find it very useful to include discontinuous signals such as square waves in our discussions, it is worthwhile to investigate the issue of convergence in a bit more detail. Specifically, there are two somewhat different classes of conditions that a periodic signal can satisfy to guarantee that it can be represented by a Fourier series. In discussing these, we will not attempt to provide a complete mathematical justification; more rigorous treatments can be found in many texts on Fourier analysis.[9]

---

[9] See, for example, R. V. Churchill, *Fourier Series and Boundary Value Problems*, 3rd ed. (New York: McGraw-Hill Book Company, 1978); W. Kaplan, *Operational Methods for Linear Systems* (Reading, MA: Addison-Wesley Publishing Company, 1962); and the book by Dym and McKean referenced in footnote 2 of this chapter.

One class of periodic signals that are representable through the Fourier series is those signals which have finite energy over a single period, i.e., signals for which

$$\int_T |x(t)|^2 \, dt < \infty. \tag{3.51}$$

When this condition is satisfied, we are guaranteed that the coefficients a_k obtained from eq. (3.39) are finite. Furthermore, let $x_N(t)$ be the approximation to $x(t)$ obtained by using these coefficients for $|k| \leq N$:

$$x_N(t) = \sum_{k=-N}^{+N} a_k e^{jk\omega_0 t}. \tag{3.52}$$

Then we are guaranteed that the energy E_N in the approximation error, as defined in eq. (3.49), converges to 0 as we add more and more terms, i.e., as $N \to \infty$. That is, if we define

$$e(t) = x(t) - \sum_{k=-\infty}^{+\infty} a_k e^{jk\omega_0 t}, \tag{3.53}$$

then

$$\int_T |e(t)|^2 \, dt = 0. \tag{3.54}$$

As we will see in an example at the end of this section, eq. (3.54) does *not* imply that the signal $x(t)$ and its Fourier series representation

$$\sum_{k=-\infty}^{+\infty} a_k e^{jk\omega_0 t} \tag{3.55}$$

are equal at every value of t. What it does say is that there is no energy in their difference.

The type of convergence guaranteed when $x(t)$ has finite energy over a single period is quite useful. In this case eq. (3.54) states that the difference between $x(t)$ and its Fourier series representation has zero energy. Since physical systems respond to signal energy, from this perspective $x(t)$ and its Fourier series representation are indistinguishable. Because most of the periodic signals that we consider do have finite energy over a single period, they have Fourier series representations. Moreover, an alternative set of conditions, developed by P. L. Dirichlet and also satisfied by essentially all of the signals with which we will be concerned, guarantees that $x(t)$ *equals* its Fourier series representation, except at isolated values of t for which $x(t)$ is discontinuous. At these values, the infinite series of eq. (3.55) converges to the average of the values on either side of the discontinuity.

The Dirichlet conditions are as follows:

Condition 1. Over any period, $x(t)$ must be *absolutely integrable*; that is,

$$\int_T |x(t)| \, dt < \infty. \tag{3.56}$$

As with square integrability, this guarantees that each coefficient a_k will be finite, since

$$|a_k| \leq \frac{1}{T}\int_T |x(t)e^{-jk\omega_0 t}|\,dt = \frac{1}{T}\int_T |x(t)|\,dt.$$

So if

$$\int_T |x(t)|\,dt < \infty,$$

then

$$|a_k| < \infty.$$

A periodic signal that violates the first Dirichlet condition is

$$x(t) = \frac{1}{t}, \quad 0 < t \leq 1;$$

that is, $x(t)$ is periodic with period 1. This signal is illustrated in Figure 3.8(a).

Condition 2. In any finite interval of time, $x(t)$ is of bounded variation; that is, there are no more than a finite number of maxima and minima during any single period of the signal.

An example of a function that meets Condition 1 but not Condition 2 is

$$x(t) = \sin\left(\frac{2\pi}{t}\right), \quad 0 < t \leq 1, \tag{3.57}$$

as illustrated in Figure 3.8(b). For this function, which is periodic with $T = 1$,

$$\int_0^1 |x(t)|\,dt < 1.$$

The function has, however, an infinite number of maxima and minima in the interval.

Condition 3. In any finite interval of time, there are only a finite number of discontinuities. Furthermore, each of these discontinuities is finite.

An example of a function that violates Condition 3 is illustrated in Figure 3.8(c). The signal, of period $T = 8$, is composed of an infinite number of sections, each of which is half the height and half the width of the previous section. Thus, the area under one period of the function is clearly less than 8. However, there are an infinite number of discontinuities in each period, thereby violating Condition 3.

As can be seen from the examples given in Figure 3.8, signals that do not satisfy the Dirichlet conditions are generally pathological in nature and consequently do not typically arise in practical contexts. For this reason, the question of the convergence of Fourier series will not play a particularly significant role in the remainder of the book. For a periodic signal that has no discontinuities, the Fourier series representation converges and equals the original signal at every value of t. For a periodic signal with a finite number of discontinuities in each period, the Fourier series representation equals the signal every-

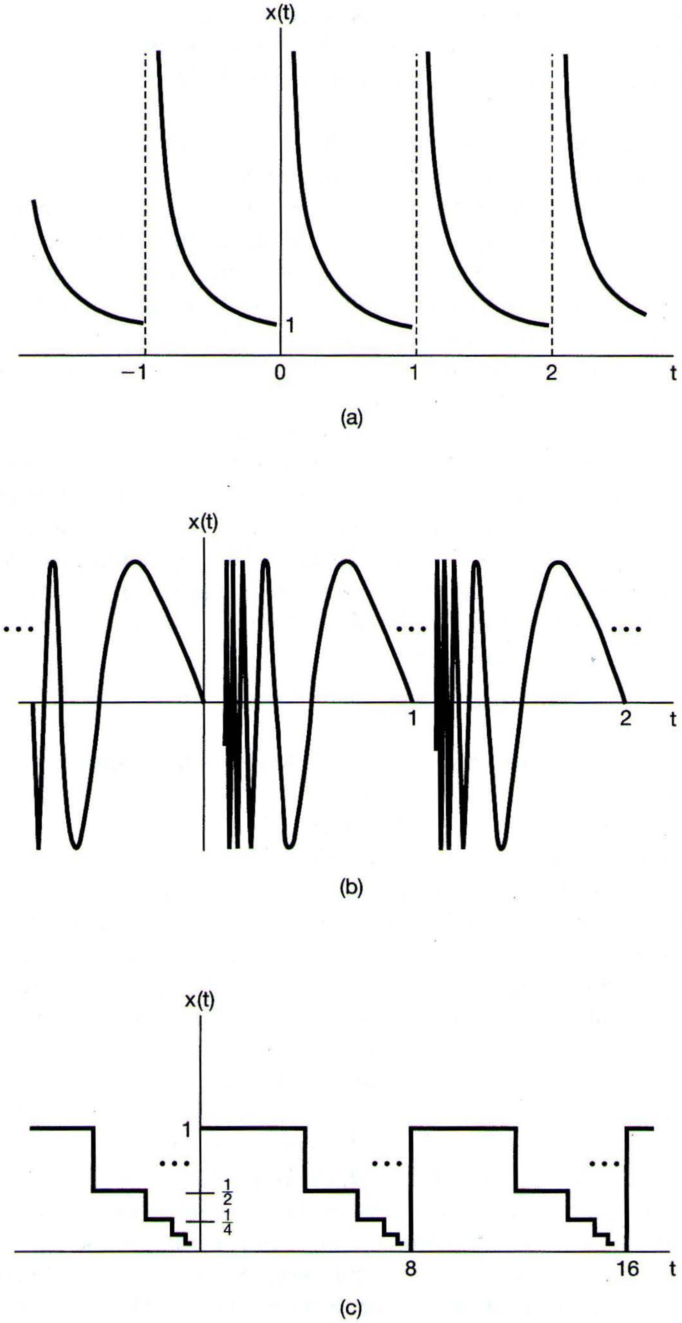


Figure 3.8 Signals that violate the Dirichlet conditions: (a) the signal $x(t) = 1/t$ for $0 < t \le 1$, a periodic signal with period 1 (this signal violates the first Dirichlet condition); (b) the periodic signal of eq. (3.57), which violates the second Dirichlet condition; (c) a signal periodic with period 8 that violates the third Dirichlet condition [for $0 \le t < 8$, the value of $x(t)$ decreases by a factor of 2 whenever the distance from t to 8 decreases by a factor of 2; that is, $x(t) = 1$, $0 \le t < 4$, $x(t) = 1/2$, $4 \le t < 6$, $x(t) = 1/4$, $6 \le t < 7$, $x(t) = 1/8$, $7 \le t < 7.5$, etc.].

where except at the isolated points of discontinuity, at which the series converges to the average value of the signal on either side of the discontinuity. In this case the difference between the original signal and its Fourier series representation contains no energy, and consequently, the two signals can be thought of as being the same for all practical purposes. Specifically, since the signals differ only at isolated points, the integrals of both signals over any interval *are* identical. For this reason, the two signals behave identically under convolution and consequently are identical from the standpoint of the analysis of LTI systems.

To gain some additional understanding of *how* the Fourier series converges for a periodic signal with discontinuities, let us return to the example of a square wave. In particular, in 1898,[10] an American physicist, Albert Michelson, constructed a harmonic analyzer, a device that, for any periodic signal $x(t)$, would compute the truncated Fourier series approximation of eq. (3.52) for values of N up to 80. Michelson tested his device on many functions, with the expected result that $x_N(t)$ looked very much like $x(t)$. However, when he tried the square wave, he obtained an important and, to him, very surprising result. Michelson was concerned about the behavior he observed and thought that his device might have had a defect. He wrote about the problem to the famous mathematical physicist Josiah Gibbs, who investigated it and reported his explanation in 1899.

What Michelson had observed is illustrated in Figure 3.9, where we have shown $x_N(t)$ for several values of N for $x(t)$, a symmetric square wave ($T = 4T_1$). In each case, the partial sum is superimposed on the original square wave. Since the square wave satisfies the Dirichlet conditions, the limit as $N \to \infty$ of $x_N(t)$ at the discontinuities should be the average value of the discontinuity. We see from the figure that this is in fact the case, since for any N, $x_N(t)$ has exactly that value at the discontinuities. Furthermore, for any other value of t, say, $t = t_1$, we are guaranteed that

$$\lim_{N\to\infty} x_N(t_1) = x(t_1).$$

Therefore, the squared error in the Fourier series representation of the square wave has zero area, as in eqs. (3.53) and (3.54).

For this example, the interesting effect that Michelson observed is that the behavior of the partial sum in the vicinity of the discontinuity exhibits ripples and that the peak amplitude of these ripples does not seem to decrease with increasing N. Gibbs showed that these are in fact the case. Specifically, for a discontinuity of unity height, the partial sum exhibits a maximum value of 1.09 (i.e., an overshoot of 9% of the height of the discontinuity), no matter how large N becomes. One must be careful to interpret this correctly, however. As stated before, for any *fixed* value of t, say, $t = t_1$, the partial sums will converge to the correct value, and at the discontinuity they will converge to one-half the sum of the values of the signal on either side of the discontinuity. However, the closer t_1 is chosen to the point of discontinuity, the larger N must be in order to reduce the error below a specified amount. Thus, as N increases, the ripples in the partial sums become compressed toward the discontinuity, but for *any* finite value of N, the peak amplitude of the ripples remains constant. This behavior has come to be known as the *Gibbs phenomenon*. The implication is that the truncated Fourier series approximation $x_N(t)$ of a discontinuous signal $x(t)$ will in general exhibit high-frequency ripples and overshoot $x(t)$ near the discontinuities. If such an approximation is used in practice, a large enough value of N should be chosen so as to guarantee that the total energy in these ripples is insignificant. In the limit, of course, we know that the energy in the approximation error vanishes and that the Fourier series representation of a discontinuous signal such as the square wave converges.

---

[10] The historical information used in this example is taken from the book by Lanczos referenced in footnote 1 of this chapter.

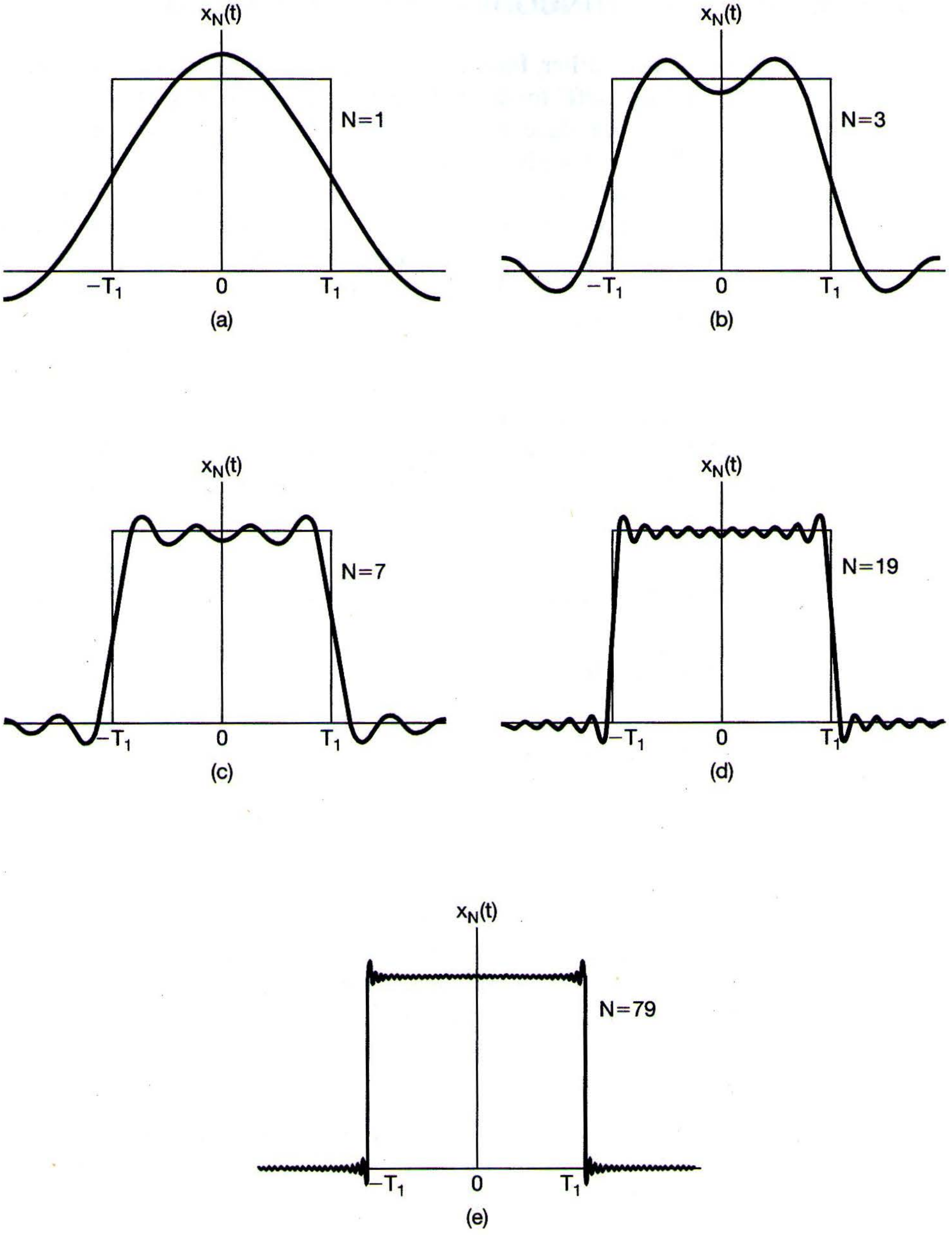


Figure 3.9 Convergence of the Fourier series representation of a square wave: an illustration of the Gibbs phenomenon. Here, we have depicted the finite series approximation $x_N(t) = \sum_{k=-N}^{N} a_k e^{jk\omega_0 t}$ for several values of N.

3.5 PROPERTIES OF CONTINUOUS-TIME FOURIER SERIES

As mentioned earlier, Fourier series representations possess a number of important properties that are useful for developing conceptual insights into such representations, and they can also help to reduce the complexity of the evaluation of the Fourier series of many signals. In Table 3.1 we have summarized these properties, several of which are considered in the problems at the end of this chapter. In Chapter 4, in which we develop the Fourier transform, we will see that most of these properties can be deduced from corresponding properties of the continuous-time Fourier transform. Consequently we limit ourselves here to the discussion of several of these properties to illustrate how they may be derived, interpreted, and used.

Throughout the following discussion of selected properties from Table 3.1, we will find it convenient to use a shorthand notation to indicate the relationship between a periodic signal and its Fourier series coefficients. Specifically, suppose that $x(t)$ is a periodic signal with period T and fundamental frequency $\omega_0 = 2\pi/T$. Then if the Fourier series coefficients of $x(t)$ are denoted by a_k, we will use the notation

$$x(t) \overset{\mathcal{FS}}{\longleftrightarrow} a_k$$

to signify the pairing of a periodic signal with its Fourier series coefficients.

3.5.1 Linearity

Let $x(t)$ and $y(t)$ denote two periodic signals with period T and which have Fourier series coefficients denoted by a_k and b_k, respectively. That is,

$$x(t) \overset{\mathcal{FS}}{\longleftrightarrow} a_k,$$
$$y(t) \overset{\mathcal{FS}}{\longleftrightarrow} b_k.$$

Since $x(t)$ and $y(t)$ have the same period T, it easily follows that any linear combination of the two signals will also be periodic with period T. Furthermore, the Fourier series coefficients c_k of the linear combination of $x(t)$ and $y(t)$, $z(t) = Ax(t) + By(t)$, are given by the same linear combination of the Fourier series coefficients for $x(t)$ and $y(t)$. That is,

$$z(t) = Ax(t) + By(t) \overset{\mathcal{FS}}{\longleftrightarrow} c_k = Aa_k + Bb_k. \tag{3.58}$$

The proof of this follows directly from the application of eq. (3.39). We also note that the linearity property is easily extended to a linear combination of an arbitrary number of signals with period T.

3.5.2 Time Shifting

When a time shift is applied to a periodic signal $x(t)$, the period T of the signal is preserved. The Fourier series coefficients b_k of the resulting signal $y(t) = x(t-t_0)$ may be expressed as

$$b_k = \frac{1}{T}\int_T x(t-t_0)e^{-jk\omega_0 t}dt. \tag{3.59}$$

Letting $\tau = t - t_0$ in the integral, and noting that the new variable τ will also range over an interval of duration T, we obtain

$$\begin{aligned}\frac{1}{T}\int_T x(\tau)e^{-jk\omega_0(\tau+t_0)}d\tau &= e^{-jk\omega_0 t_0}\frac{1}{T}\int_T x(\tau)e^{-jk\omega_0\tau}d\tau \\ &= e^{-jk\omega_0 t_0}a_k = e^{-jk(2\pi/T)t_0}a_k,\end{aligned} \tag{3.60}$$

where a_k is the kth Fourier series coefficient of $x(t)$. That is, if

$$x(t) \overset{\mathfrak{F}S}{\longleftrightarrow} a_k,$$

then

$$x(t-t_0) \overset{\mathfrak{F}S}{\longleftrightarrow} e^{-jk\omega_0 t_0}a_k = e^{-jk(2\pi/T)t_0}a_k.$$

One consequence of this property is that, when a periodic signal is shifted in time, the *magnitudes* of its Fourier series coefficients remain unaltered. That is, $|b_k| = |a_k|$.

3.5.3 Time Reversal

The period T of a periodic signal $x(t)$ also remains unchanged when the signal undergoes time reversal. To determine the Fourier series coefficients of $y(t) = x(-t)$, let us consider the effect of time reversal on the synthesis equation (3.38):

$$x(-t) = \sum_{k=-\infty}^{\infty} a_k e^{-jk2\pi t/T}. \tag{3.61}$$

Making the substitution $k = -m$, we obtain

$$y(t) = x(-t) = \sum_{m=-\infty}^{\infty} a_{-m}e^{jm2\pi t/T}. \tag{3.62}$$

We observe that the right-hand side of this equation has the form of a Fourier series synthesis equation for $x(-t)$, where the Fourier series coefficients b_k are

$$b_k = a_{-k}. \tag{3.63}$$

That is, if

$$x(t) \overset{\mathfrak{F}S}{\longleftrightarrow} a_k,$$

then

$$x(-t) \overset{\mathfrak{F}S}{\longleftrightarrow} a_{-k}.$$

In other words time reversal applied to a continuous-time signal results in a time reversal of the corresponding sequence of Fourier series coefficients. An interesting consequence of the time-reversal property is that if $x(t)$ is even—that is, if $x(-t) = x(t)$—then its Fourier series coefficients are also even—i.e., $a_{-k} = a_k$. Similarly, if $x(t)$ is odd, so that $x(-t) = -x(t)$, then so are its Fourier series coefficients—i.e., $a_{-k} = -a_k$.

3.5.4 Time Scaling

Time scaling is an operation that in general changes the period of the underlying signal. Specifically, if $x(t)$ is periodic with period T and fundamental frequency $\omega_0 = 2\pi/T$, then $x(\alpha t)$, where α is a positive real number, is periodic with period T/α and fundamental frequency $\alpha\omega_0$. Since the time-scaling operation applies directly to each of the harmonic components of $x(t)$, we may easily conclude that the Fourier coefficients for each of those components remain the same. That is, if $x(t)$ has the Fourier series representation in eq. (3.38), then

$$x(\alpha t) = \sum_{k=-\infty}^{+\infty} a_k e^{jk(\alpha\omega_0)t}$$

is the Fourier series representation of $x(\alpha t)$. We emphasize that, while the Fourier coefficients have not changed, the Fourier series representation *has* changed because of the change in the fundamental frequency.

3.5.5 Multiplication

Suppose that $x(t)$ and $y(t)$ are both periodic with period T and that

$$x(t) \overset{\mathfrak{FS}}{\longleftrightarrow} a_k,$$

$$y(t) \overset{\mathfrak{FS}}{\longleftrightarrow} b_k.$$

Since the product $x(t)y(t)$ is also periodic with period T, we can expand it in a Fourier series with Fourier series coefficients h_k expressed in terms of those for $x(t)$ and $y(t)$. The result is

$$x(t)y(t) \overset{\mathfrak{FS}}{\longleftrightarrow} h_k = \sum_{l=-\infty}^{\infty} a_l b_{k-l}. \qquad (3.64)$$

One way to derive this relationship (see Problem 3.46) is to multiply the Fourier series representations of $x(t)$ and $y(t)$ and to note that the kth harmonic component in the product will have a coefficient which is the sum of terms of the form $a_l b_{k-l}$. Observe that the sum on the right-hand side of eq. (3.64) may be interpreted as the discrete-time convolution of the sequence representing the Fourier coefficients of $x(t)$ and the sequence representing the Fourier coefficients of $y(t)$.

3.5.6 Conjugation and Conjugate Symmetry

Taking the complex conjugate of a periodic signal $x(t)$ has the effect of complex conjugation *and* time reversal on the corresponding Fourier series coefficients. That is, if

$$x(t) \overset{\mathfrak{FS}}{\longleftrightarrow} a_k,$$

then

$$x^*(t) \overset{\mathfrak{FS}}{\longleftrightarrow} a^*_{-k}. \qquad (3.65)$$

This property is easily proved by applying complex conjugation to both sides of eq. (3.38) and replacing the summation variable k by its negative.

Some interesting consequences of this property may be derived for $x(t)$ real—that is, when $x(t) = x^*(t)$. In particular, in this case, we see from eq. (3.65) that the Fourier series coefficients will be *conjugate symmetric*, i.e.,

$$a_{-k} = a_k^*, \tag{3.66}$$

as we previously saw in eq. (3.29). This in turn implies various symmetry properties (listed in Table 3.1) for the magnitudes, phases, real parts, and imaginary parts of the Fourier series coefficients of real signals. For example, from eq. (3.66), we see that if $x(t)$ is real, then a_0 is real and

$$|a_k| = |a_{-k}|.$$

Also, if $x(t)$ is real and even, then, from Section 3.5.3, $a_k = a_{-k}$. However, from eq. (3.66) we see that $a_k^* = a_{-k}$, so that $a_k = a_k^*$. That is, if $x(t)$ is real and even, then so are its Fourier series coefficients. Similarly, it can be shown that if $x(t)$ is real and odd, then its Fourier series coefficients are purely imaginary and odd. Thus, for example, $a_0 = 0$ if $x(t)$ is real and odd. This and the other symmetry properties of the Fourier series are examined further in Problem 3.42.

3.5.7 Parseval's Relation for Continuous-Time Periodic Signals

As shown in Problem 3.46, Parseval's relation for continuous-time periodic signals is

$$\frac{1}{T}\int_T |x(t)|^2 dt = \sum_{k=-\infty}^{+\infty} |a_k|^2, \tag{3.67}$$

where the a_k are the Fourier series coefficients of $x(t)$ and T is the period of the signal.

Note that the left-hand side of eq. (3.67) is the average power (i.e., energy per unit time) in one period of the periodic signal $x(t)$. Also,

$$\frac{1}{T}\int_T \left|a_k e^{jk\omega_0 t}\right|^2 dt = \frac{1}{T}\int_T |a_k|^2 dt = |a_k|^2, \tag{3.68}$$

so that $|a_k|^2$ is the average power in the kth harmonic component of $x(t)$. Thus, what Parseval's relation states is that the total average power in a periodic signal equals the sum of the average powers in all of its harmonic components.

3.5.8 Summary of Properties of the Continuous-Time Fourier Series

In Table 3.1, we summarize these and other important properties of continuous-time Fourier series.

3.5.9 Examples

Fourier series properties, such as those listed in Table 3.1, may be used to circumvent some of the algebra involved in determining the Fourier coefficients of a given signal. In the next

TABLE 3.1 PROPERTIES OF CONTINUOUS-TIME FOURIER SERIES

Property	Section	Periodic Signal	Fourier Series Coefficients
		$\left.\begin{matrix}x(t)\\ y(t)\end{matrix}\right\}$ Periodic with period T and fundamental frequency $\omega_0 = 2\pi/T$	a_k b_k
Linearity	3.5.1	$Ax(t) + By(t)$	$Aa_k + Bb_k$
Time Shifting	3.5.2	$x(t - t_0)$	$a_k e^{-jk\omega_0 t_0} = a_k e^{-jk(2\pi/T)t_0}$
Frequency Shifting		$e^{jM\omega_0 t}x(t) = e^{jM(2\pi/T)t}x(t)$	a_{k-M}
Conjugation	3.5.6	$x^*(t)$	a^*_{-k}
Time Reversal	3.5.3	$x(-t)$	a_{-k}
Time Scaling	3.5.4	$x(\alpha t)$, $\alpha > 0$ (periodic with period T/α)	a_k
Periodic Convolution		$\int_T x(\tau)y(t-\tau)d\tau$	$Ta_k b_k$
Multiplication	3.5.5	$x(t)y(t)$	$\sum_{l=-\infty}^{+\infty} a_l b_{k-l}$
Differentiation		$\frac{dx(t)}{dt}$	$jk\omega_0 a_k = jk\frac{2\pi}{T}a_k$
Integration		$\int_{-\infty}^{t} x(t)\,dt$ (finite valued and periodic only if $a_0 = 0$)	$\left(\frac{1}{jk\omega_0}\right)a_k = \left(\frac{1}{jk(2\pi/T)}\right)a_k$
Conjugate Symmetry for Real Signals	3.5.6	$x(t)$ real	$a_k = a^*_{-k}$ $\mathcal{R}e\{a_k\} = \mathcal{R}e\{a_{-k}\}$ $\mathcal{I}m\{a_k\} = -\mathcal{I}m\{a_{-k}\}$ $\lvert a_k\rvert = \lvert a_{-k}\rvert$ $\sphericalangle a_k = -\sphericalangle a_{-k}$
Real and Even Signals	3.5.6	$x(t)$ real and even	a_k real and even
Real and Odd Signals	3.5.6	$x(t)$ real and odd	a_k purely imaginary and odd
Even-Odd Decomposition of Real Signals		$x_e(t) = \mathcal{E}v\{x(t)\}$ [$x(t)$ real]	$\mathcal{R}e\{a_k\}$
		$x_o(t) = \mathcal{O}d\{x(t)\}$ [$x(t)$ real]	$j\mathcal{I}m\{a_k\}$

Parseval's Relation for Periodic Signals

$$\frac{1}{T}\int_T |x(t)|^2 dt = \sum_{k=-\infty}^{+\infty} |a_k|^2$$

three examples, we illustrate this. The last example in this section then demonstrates how properties of a signal can be used to characterize the signal in great detail.

Example 3.6

Consider the signal $g(t)$ with a fundamental period of 4, shown in Figure 3.10. We could determine the Fourier series representation of $g(t)$ directly from the analysis equation (3.39). Instead, we will use the relationship of $g(t)$ to the symmetric periodic square wave $x(t)$ in Example 3.5. Referring to that example, we see that, with $T = 4$ and $T_1 = 1$,

$$g(t) = x(t-1) - 1/2. \tag{3.69}$$

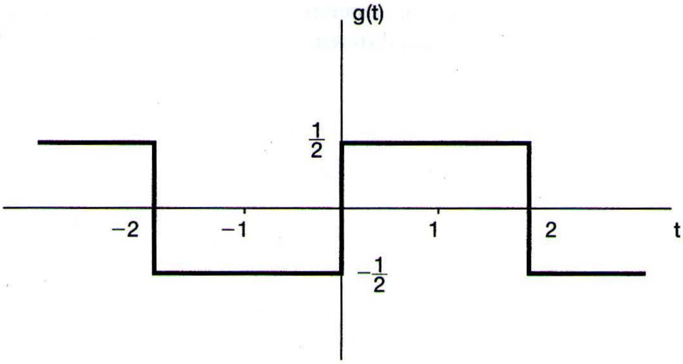


Figure 3.10 Periodic signal for Example 3.6.

The time-shift property in Table 3.1 indicates that, if the Fourier Series coefficients of $x(t)$ are denoted by a_k, the Fourier coefficients of $x(t-1)$ may be expressed as

$$b_k = a_k e^{-jk\pi/2}. \tag{3.70}$$

The Fourier coefficients of the *dc offset* in $g(t)$—i.e., the term $-1/2$ on the right-hand side of eq. (3.69)—are given by

$$c_k = \begin{cases} 0, & \text{for } k \neq 0 \\ -\frac{1}{2}, & \text{for } k = 0 \end{cases}. \tag{3.71}$$

Applying the linearity property in Table 3.1, we conclude that the coefficients for $g(t)$ may be expressed as

$$d_k = \begin{cases} a_k e^{-jk\pi/2}, & \text{for } k \neq 0 \\ a_0 - \frac{1}{2}, & \text{for } k = 0 \end{cases},$$

where each a_k may now be replaced by the corresponding expression from eqs. (3.45) and (3.46), yielding

$$d_k = \begin{cases} \frac{\sin(\pi k/2)}{k\pi} e^{-jk\pi/2}, & \text{for } k \neq 0 \\ 0, & \text{for } k = 0 \end{cases}. \tag{3.72}$$

Example 3.7

Consider the triangular wave signal $x(t)$ with period $T = 4$ and fundamental frequency $\omega_0 = \pi/2$ shown in Figure 3.11. The derivative of this signal is the signal $g(t)$ in Exam-

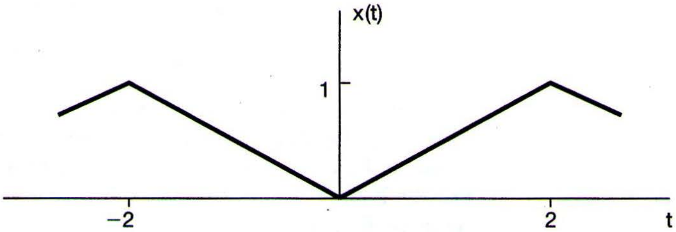


Figure 3.11 Triangular wave signal in Example 3.7.

ple 3.6. Denoting the Fourier coefficients of $g(t)$ by d_k and those of $x(t)$ by e_k, we see that the differentiation property in Table 3.1 indicates that

$$d_k = jk(\pi/2)e_k. \tag{3.73}$$

This equation can be used to express e_k in terms of d_k, except when $k = 0$. Specifically, from eq. (3.72),

$$e_k = \frac{2d_k}{jk\pi} = \frac{2\sin(\pi k/2)}{j(k\pi)^2}e^{-jk\pi/2}, \quad k \neq 0. \tag{3.74}$$

For $k = 0$, e_0 can be determined by finding the area under one period of $x(t)$ and dividing by the length of the period:

$$e_0 = \frac{1}{2}.$$

Example 3.8

Let us examine some properties of the Fourier series representation of a periodic train of impulses, or *impulse train.* This signal and its representation in terms of complex exponentials will play an important role when we discuss the topic of sampling in Chapter 7. The impulse train with period T may be expressed as

$$x(t) = \sum_{k=-\infty}^{\infty} \delta(t - kT); \tag{3.75}$$

it is illustrated in Figure 3.12(a). To determine the Fourier series coefficients a_k, we use eq. (3.39) and select the interval of integration to be $-T/2 \leq t \leq T/2$, avoiding the placement of impulses at the integration limits. Within this interval, $x(t)$ is the same as $\delta(t)$, and it follows that

$$a_k = \frac{1}{T}\int_{-T/2}^{T/2} \delta(t)e^{-jk2\pi t/T}dt = \frac{1}{T}. \tag{3.76}$$

In other words, all the Fourier series coefficients of the impulse train are identical. These coefficients are also real valued and even (with respect to the index k). This is to be expected, since, according to Table 3.1, any real and even signal (such as our impulse train) should have real and even Fourier coefficients.

The impulse train also has a straightforward relationship to square-wave signals such as $g(t)$ in Figure 3.6, which we repeat in Figure 3.12(b). The derivative of $g(t)$ is the signal $q(t)$ illustrated in Figure 3.12(c). We may interpret $q(t)$ as the difference of two shifted versions of the impulse train $x(t)$. That is,

$$q(t) = x(t + T_1) - x(t - T_1). \tag{3.77}$$

Using the properties of Fourier series, we can now compute the Fourier series coefficients of $q(t)$ and $g(t)$ without any further direct evaluation of the Fourier series analysis equation. First, from the time-shifting and linearity properties, we see from eq. (3.77) that the Fourier series coefficients b_k of $q(t)$ may be expressed in terms of the Fourier series coefficients a_k of $x(t)$; that is,

$$b_k = e^{jk\omega_0 T_1}a_k - e^{-jk\omega_0 T_1}a_k,$$

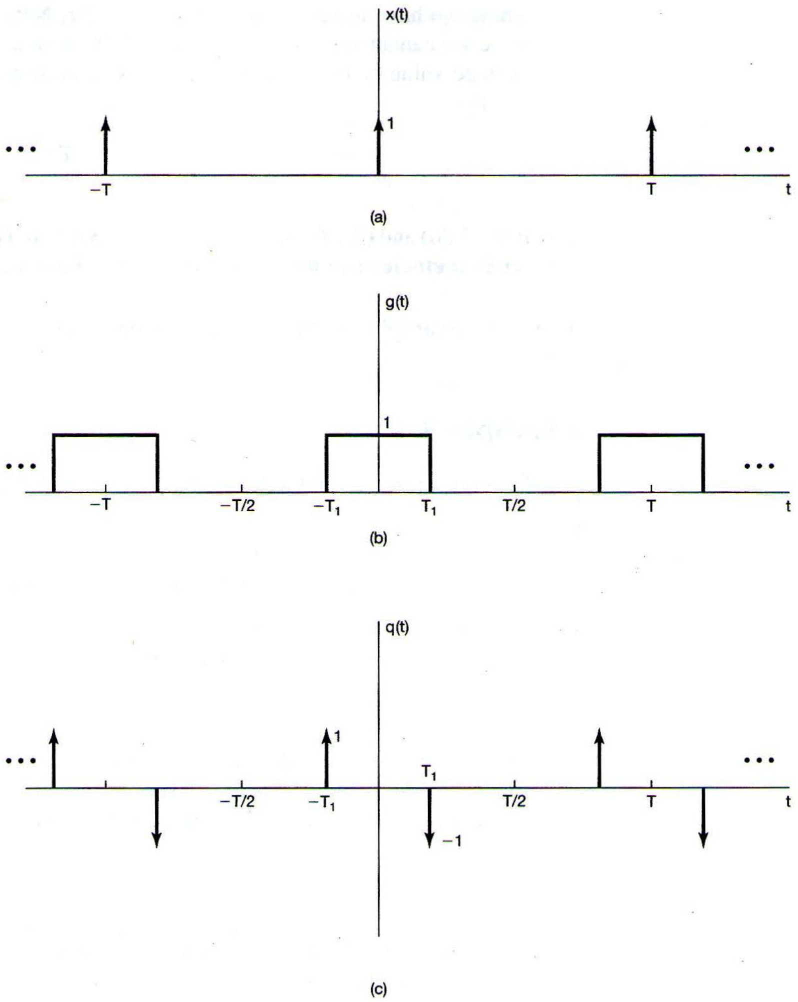


Figure 3.12 (a) Periodic train of impulses; (b) periodic square wave; (c) derivative of the periodic square wave in (b).

where $\omega_0 = 2\pi/T$. Using eq. (3.76), we then have

$$b_k = \frac{1}{T}[e^{jk\omega_0 T_1} - e^{-jk\omega_0 T_1}] = \frac{2j\sin(k\omega_0 T_1)}{T}.$$

Finally, since $q(t)$ is the derivative of $g(t)$, we can use the differentiation property in Table 3.1 to write

$$b_k = jk\omega_0 c_k, \tag{3.78}$$

where the c_k are the Fourier series coefficients of $g(t)$. Thus,

$$c_k = \frac{b_k}{jk\omega_0} = \frac{2j\sin(k\omega_0 T_1)}{jk\omega_0 T} = \frac{\sin(k\omega_0 T_1)}{k\pi}, \quad k \neq 0, \tag{3.79}$$

where we have used the fact that $\omega_0 T = 2\pi$. Note that eq. (3.79) is valid for $k \neq 0$, since we cannot solve for c_0 from eq. (3.78) with $k = 0$. However, since c_0 is just the average value of $g(t)$ over one period, we can determine it by inspection from Figure 3.12(b):

$$c_0 = \frac{2T_1}{T}. \tag{3.80}$$

Eqs. (3.80) and (3.79) are identical to eqs. (3.42) and (3.44), respectively, for the Fourier series coefficients of the square wave derived in Example 3.5.

The next example is chosen to illustrate the use of many of the properties in Table 3.1.

Example 3.9

Suppose we are given the following facts about a signal $x(t)$:

1. $x(t)$ is a real signal.
2. $x(t)$ is periodic with period $T = 4$, and it has Fourier series coefficients a_k.
3. $a_k = 0$ for $|k| > 1$.
4. The signal with Fourier coefficients $b_k = e^{-j\pi k/2}a_{-k}$ is odd.
5. $\frac{1}{4}\int_4 |x(t)|^2 dt = 1/2$.

Let us show that this information is sufficient to determine the signal $x(t)$ to within a sign factor. According to Fact 3, $x(t)$ has at most three nonzero Fourier series coefficients a_k: a_0, a_1, and a_{-1}. Then, since $x(t)$ has fundamental frequency $\omega_0 = 2\pi/4 = \pi/2$, it follows that

$$x(t) = a_0 + a_1 e^{j\pi t/2} + a_{-1}e^{-j\pi t/2}.$$

Since $x(t)$ is real (Fact 1), we can use the symmetry properties in Table 3.1 to conclude that a_0 is real and $a_1 = a*_{-1}$. Consequently,

$$x(t) = a_0 + a_1 e^{j\pi t/2} + (a_1 e^{j\pi t/2})^* = a_0 + 2\mathcal{R}e\{a_1 e^{j\pi t/2}\}. \tag{3.81}$$

Let us now determine the signal corresponding to the Fourier coefficients b_k given in Fact 4. Using the time-reversal property from Table 3.1, we note that a_{-k} corresponds to the signal $x(-t)$. Also, the time-shift property in the table indicates that multiplication of the kth Fourier coefficient by $e^{-jk\pi/2} = e^{-jk\omega_0}$ corresponds to the underlying signal being shifted by 1 to the right (i.e., having t replaced by $t - 1$). We conclude that the coefficients b_k correspond to the signal $x(-(t-1)) = x(-t+1)$, which, according to Fact 4, must be odd. Since $x(t)$ is real, $x(-t+1)$ must also be real. From Table 3.1, it then follows that the Fourier coefficients of $x(-t+1)$ must be purely imaginary and odd. Thus, $b_0 = 0$ and $b_{-1} = -b_1$. Since time-reversal and time-shift operations cannot change the average power per period, Fact 5 holds even if $x(t)$ is replaced by $x(-t+1)$. That is,

$$\frac{1}{4}\int_4 |x(-t+1)|^2 dt = 1/2. \tag{3.82}$$

We can now use Parseval's relation to conclude that

$$|b_1|^2 + |b_{-1}|^2 = 1/2. \tag{3.83}$$

Substituting $b_1 = -b_{-1}$ in this equation, we obtain $|b_1| = 1/2$. Since b_1 is also known to be purely imaginary, it must be either $j/2$ or $-j/2$.

Now we can translate these conditions on b_0 and b_1 into equivalent statements on a_0 and a_1. First, since $b_0 = 0$, Fact 4 implies that $a_0 = 0$. With $k = 1$, this condition implies that $a_1 = e^{-j\pi/2}b_{-1} = -jb_{-1} = jb_1$. Thus, if we take $b_1 = j/2$, then $a_1 = -1/2$, and therefore, from eq. (3.81), $x(t) = -\cos(\pi t/2)$. Alternatively, if we take $b_1 = -j/2$, then $a_1 = 1/2$, and therefore, $x(t) = \cos(\pi t/2)$.

3.6 FOURIER SERIES REPRESENTATION OF DISCRETE-TIME PERIODIC SIGNALS

In this section, we consider the Fourier series representation of discrete-time periodic signals. While the discussion closely parallels that of Section 3.3, there are some important differences. In particular, the Fourier series representation of a discrete-time periodic signal is a *finite* series, as opposed to the infinite series representation required for continuous-time periodic signals. As a consequence, there are no mathematical issues of convergence such as those discussed in Section 3.4.

3.6.1 Linear Combinations of Harmonically Related Complex Exponentials

As defined in Chapter 1, a discrete-time signal $x[n]$ is periodic with period N if

$$x[n] = x[n + N]. \tag{3.84}$$

The fundamental period is the smallest positive integer N for which eq. (3.84) holds, and $\omega_0 = 2\pi/N$ is the fundamental frequency. For example, the complex exponential $e^{j(2\pi/N)n}$ is periodic with period N. Furthermore, the set of all discrete-time complex exponential signals that are periodic with period N is given by

$$\phi_k[n] = e^{jk\omega_0 n} = e^{jk(2\pi/N)n}, \; k = 0, \pm 1, \pm 2, \ldots. \tag{3.85}$$

All of these signals have fundamental frequencies that are multiples of $2\pi/N$ and thus are harmonically related.

As mentioned in Section 1.3.3, there are only N distinct signals in the set given by eq. (3.85). This is a consequence of the fact that discrete-time complex exponentials which differ in frequency by a multiple of 2π are identical. Specifically, $\phi_0[n] = \phi_N[n]$, $\phi_1[n] = \phi_{N+1}[n]$, and, in general,

$$\phi_k[n] = \phi_{k+rN}[n]. \tag{3.86}$$

That is, when k is changed by any integer multiple of N, the identical sequence is generated. This differs from the situation in continuous time in which the signals $\phi_k(t)$ defined in eq. (3.24) are all different from one another.

We now wish to consider the representation of more general periodic sequences in terms of linear combinations of the sequences $\phi_k[n]$ in eq. (3.85). Such a linear combination has the form

$$x[n] = \sum_k a_k \phi_k[n] = \sum_k a_k e^{jk\omega_0 n} = \sum_k a_k e^{jk(2\pi/N)n}. \tag{3.87}$$

Since the sequences $\phi_k[n]$ are distinct only over a range of N successive values of k, the summation in eq. (3.87) need only include terms over this range. Thus, the summation is on k, as k varies over a range of N successive integers, beginning with any value of k. We indicate this by expressing the limits of the summation as $k = \langle N \rangle$. That is,

$$x[n] = \sum_{k=\langle N \rangle} a_k \phi_k[n] = \sum_{k=\langle N \rangle} a_k e^{jk\omega_0 n} = \sum_{k=\langle N \rangle} a_k e^{jk(2\pi/N)n}. \tag{3.88}$$

For example, k could take on the values $k = 0, 1, \ldots, N-1$, or $k = 3, 4, \ldots, N+2$. In either case, by virtue of eq. (3.86), exactly the same set of complex exponential sequences appears in the summation on the right-hand side of eq. (3.88). Equation (3.88) is referred to as the *discrete-time Fourier series* and the coefficients a_k as the *Fourier series coefficients*.

3.6.2 Determination of the Fourier Series Representation of a Periodic Signal

Suppose now that we are given a sequence $x[n]$ that is periodic with fundamental period N. We would like to determine whether a representation of $x[n]$ in the form of eq. (3.88) exists and, if so, what the values of the coefficients a_k are. This question can be phrased in terms of finding a solution to a set of linear equations. Specifically, if we evaluate eq. (3.88) for N successive values of n corresponding to one period of $x[n]$, we obtain

$$\begin{aligned}
x[0] &= \sum_{k=\langle N \rangle} a_k, \\
x[1] &= \sum_{k=\langle N \rangle} a_k e^{j2\pi k/N}, \\
&\vdots \\
x[N-1] &= \sum_{k=\langle N \rangle} a_k e^{j2\pi k(N-1)/N}.
\end{aligned} \tag{3.89}$$

Thus, eq. (3.89) represents a set of N linear equations for the N unknown coefficients a_k as k ranges over a set of N successive integers. It can be shown that this set of equations is linearly independent and consequently can be solved to obtain the coefficients a_k in terms of the given values of $x[n]$. In Problem 3.32, we consider an example in which the Fourier series coefficients are obtained by explicitly solving the set of N equations given in eq. (3.89). However, by following steps parallel to those used in continuous time, it is possible to obtain a closed-form expression for the coefficients a_k in terms of the values of the sequence $x[n]$.

The basis for this result is the fact, shown in Problem 3.54, that

$$\sum_{n=\langle N\rangle} e^{jk(2\pi/N)n} = \begin{cases} N, & k = 0, \pm N, \pm 2N, \ldots \\ 0, & \text{otherwise} \end{cases}. \tag{3.90}$$

Equation (3.90) states that the sum over one period of the values of a periodic complex exponential is zero, unless that complex exponential is a constant.

Now consider the Fourier series representation of eq. (3.88). Multiplying both sides by $e^{-jr(2\pi/N)n}$ and summing over N terms, we obtain

$$\sum_{n=\langle N\rangle} x[n]e^{-jr(2\pi/N)n} = \sum_{n=\langle N\rangle}\sum_{k=\langle N\rangle} a_k e^{j(k-r)(2\pi/N)n}. \tag{3.91}$$

Interchanging the order of summation on the right-hand side, we have

$$\sum_{n=\langle N\rangle} x[n]e^{-jr(2\pi/N)n} = \sum_{k=\langle N\rangle} a_k \sum_{n=\langle N\rangle} e^{j(k-r)(2\pi/N)n}. \tag{3.92}$$

From the identity in eq. (3.90), the innermost sum on n on the right-hand side of eq. (3.92) is zero, unless $k - r$ is zero or an integer multiple of N. Therefore, if we choose values for r over the same range as that over which k varies in the outer summation, the innermost sum on the right-hand side of eq. (3.92) equals N if $k = r$ and 0 if $k \neq r$. The right-hand side of eq. (3.92) then reduces to Na_r, and we have

$$a_r = \frac{1}{N}\sum_{n=\langle N\rangle} x[n]e^{-jr(2\pi/N)n}. \tag{3.93}$$

This provides a closed-form expression for obtaining the Fourier series coefficients, and we have the *discrete-time Fourier series pair*:

$$\boxed{\begin{aligned} x[n] &= \sum_{k=\langle N\rangle} a_k e^{jk\omega_0 n} = \sum_{k=\langle N\rangle} a_k e^{jk(2\pi/N)n}, && (3.94)\\ a_k &= \frac{1}{N}\sum_{n=\langle N\rangle} x[n]e^{-jk\omega_0 n} = \frac{1}{N}\sum_{n=\langle N\rangle} x[n]e^{-jk(2\pi/N)n}. && (3.95) \end{aligned}}$$

These equations play the same role for discrete-time periodic signals that eqs. (3.38) and (3.39) play for continuous-time periodic signals, with eq. (3.94) the *synthesis* equation and eq. (3.95) the *analysis* equation. As in continuous time, the discrete-time Fourier series coefficients a_k are often referred to as the *spectral coefficients* of $x[n]$. These coefficients specify a decomposition of $x[n]$ into a sum of N harmonically related complex exponentials.

Referring to eq. (3.88), we see that if we take k in the range from 0 to $N - 1$, we have

$$x[n] = a_0\phi_0[n] + a_1\phi_1[n] + \ldots + a_{N-1}\phi_{N-1}[n]. \tag{3.96}$$

Similarly, if k ranges from 1 to N, we obtain

$$x[n] = a_1\phi_1[n] + a_2\phi_2[n] + \ldots + a_N\phi_N[n]. \tag{3.97}$$

From eq. (3.86), $\phi_0[n] = \phi_N[n]$, and therefore, upon comparing eqs. (3.96) and (3.97), we conclude that $a_0 = a_N$. Similarly, by letting k range over any set of N consecutive integers and using eq. (3.86), we can conclude that

$$a_k = a_{k+N}. \tag{3.98}$$

That is, if we consider more than N sequential values of k, the values a_k repeat periodically with period N. It is important that this fact be interpreted carefully. In particular, since there are only N distinct complex exponentials that are periodic with period N, the discrete-time Fourier series representation is a finite series with N terms. Therefore, if we fix the N consecutive values of k over which we define the Fourier series in eq. (3.94), we will obtain a set of exactly N Fourier coefficients from eq. (3.95). On the other hand, at times it will be convenient to use different sets of N values of k, and consequently, it is useful to regard eq. (3.94) as a sum over any *arbitrary* set of N successive values of k. For this reason, it is sometimes convenient to think of a_k as a sequence defined for all values of k, but where only N successive elements in the sequence will be used in the Fourier series representation. Furthermore, since the $\phi_k[n]$ repeat periodically with period N as we vary k [eq. (3.86)], so must the a_k [eq. (3.98)]. This viewpoint is illustrated in the next example.

Example 3.10

Consider the signal

$$x[n] = \sin \omega_0 n, \tag{3.99}$$

which is the discrete-time counterpart of the signal $x(t) = \sin \omega_0 t$ of Example 3.3. $x[n]$ is periodic only if $2\pi/\omega_0$ is an integer or a ratio of integers. For the case when $2\pi/\omega_0$ is an integer N, that is, when

$$\omega_0 = \frac{2\pi}{N},$$

$x[n]$ is periodic with fundamental period N, and we obtain a result that is exactly analogous to the continuous-time case. Expanding the signal as a sum of two complex exponentials, we get

$$x[n] = \frac{1}{2j}e^{j(2\pi/N)n} - \frac{1}{2j}e^{-j(2\pi/N)n}. \tag{3.100}$$

Comparing eq. (3.100) with eq. (3.94), we see by inspection that

$$a_1 = \frac{1}{2j}, \quad a_{-1} = -\frac{1}{2j}, \tag{3.101}$$

and the remaining coefficients over the interval of summation are zero. As described previously, these coefficients repeat with period N; thus, a_{N+1} is also equal to $(1/2j)$ and a_{N-1} equals $(-1/2j)$. The Fourier series coefficients for this example with $N = 5$ are illustrated in Figure 3.13. The fact that they repeat periodically is indicated. However, only one period is utilized in the synthesis equation (3.94).

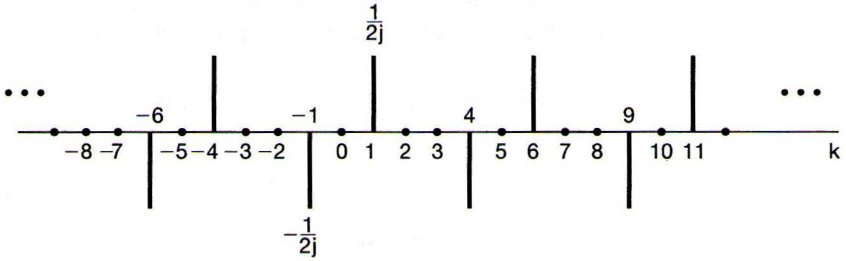


Figure 3.13 Fourier coefficients for $x[n] = \sin(2\pi/5)n$.

Consider now the case when $2\pi/\omega_0$ is a ratio of integers—that is, when

$$\omega_0 = \frac{2\pi M}{N}.$$

Assuming that M and N do not have any common factors, $x[n]$ has a fundamental period of N. Again expanding $x[n]$ as a sum of two complex exponentials, we have

$$x[n] = \frac{1}{2j}e^{jM(2\pi/N)n} - \frac{1}{2j}e^{-jM(2\pi/N)n},$$

from which we can determine by inspection that $a_M = (1/2j)$, $a_{-M} = (-1/2j)$, and the remaining coefficients over one period of length N are zero. The Fourier coefficients for this example with $M = 3$ and $N = 5$ are depicted in Figure 3.14. Again, we have indicated the periodicity of the coefficients. For example, for $N = 5$, $a_2 = a_{-3}$, which in our example equals $(-1/2j)$. Note, however, that over any period of length 5 there are only two nonzero Fourier coefficients, and therefore there are only two nonzero terms in the synthesis equation.

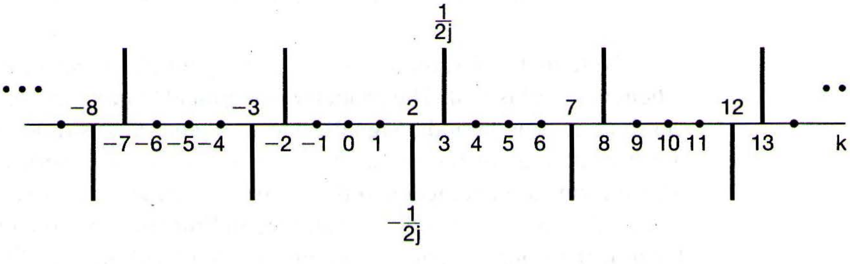
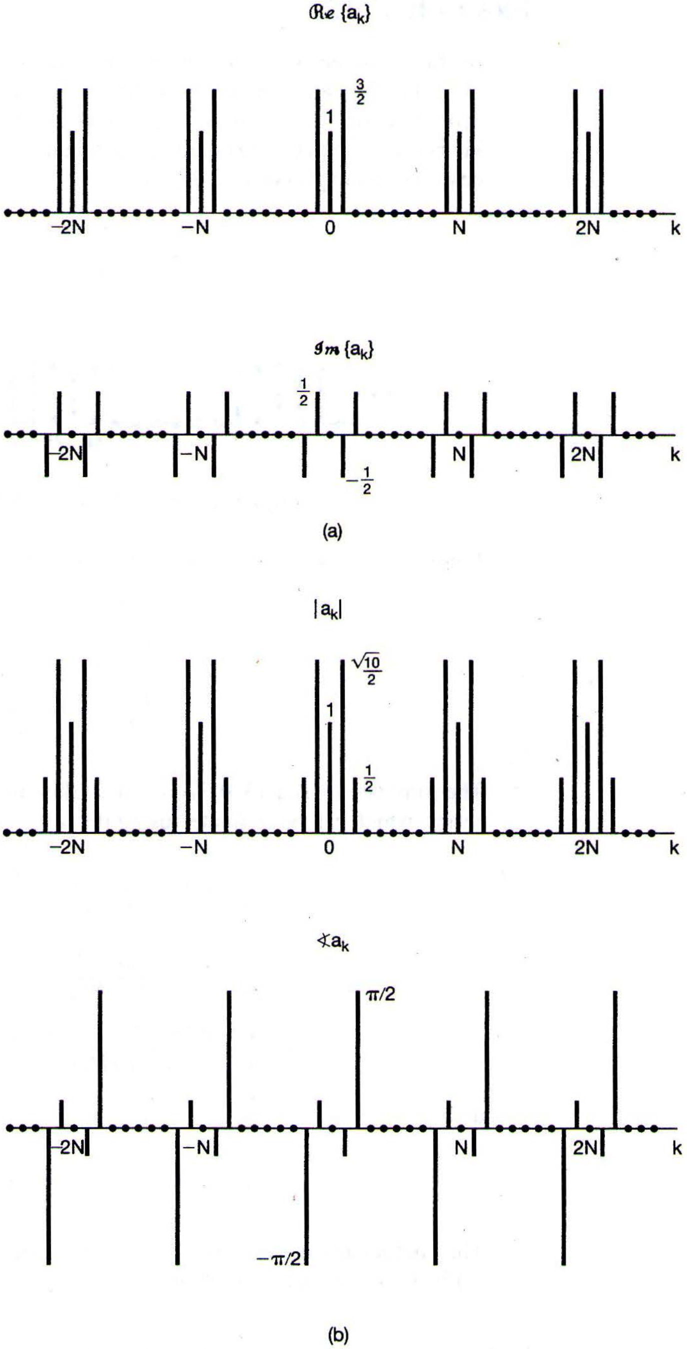


Figure 3.14 Fourier coefficients for $x[n] = \sin 3(2\pi/5)n$.



(a)

(b)

Figure 3.15 (a) Real and imaginary parts of the Fourier series coefficients in Example 3.11; (b) magnitude and phase of the same coefficients.

Example 3.12

In this example, we consider the discrete-time periodic square wave shown in Figure 3.16. We can evaluate the Fourier series for this signal using eq. (3.95). Because $x[n] = 1$ for $-N_1 \leq n \leq N_1$, it is particularly convenient to choose the length-N interval of summation in eq. (3.95) so that it includes the range $-N_1 \leq n \leq N_1$. In this case, we can express eq. (3.95) as

$$a_k = \frac{1}{N} \sum_{n=-N_1}^{N_1} e^{-jk(2\pi/N)n}. \tag{3.102}$$

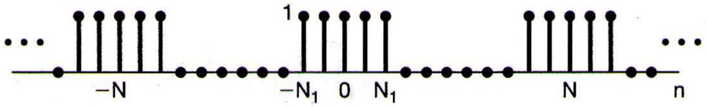


Figure 3.16 Discrete-time periodic square wave.

Letting $m = n + N_1$, we observe that eq. (3.102) becomes

$$\begin{aligned} a_k &= \frac{1}{N} \sum_{m=0}^{2N_1} e^{-jk(2\pi/N)(m-N_1)} \\ &= \frac{1}{N} e^{jk(2\pi/N)N_1} \sum_{m=0}^{2N_1} e^{-jk(2\pi/N)m}. \end{aligned} \tag{3.103}$$

The summation in eq. (3.103) consists of the sum of the first $2N_1+1$ terms in a geometric series, which can be evaluated using the result of Problem 1.54. This yields

$$\begin{aligned} a_k &= \frac{1}{N} e^{jk(2\pi/N)N_1} \left(\frac{1 - e^{-jk2\pi(2N_1+1)/N}}{1 - e^{-jk(2\pi/N)}} \right) \\ &= \frac{1}{N} \frac{e^{-jk(2\pi/2N)}[e^{jk2\pi(N_1+1/2)/N} - e^{-jk2\pi(N_1+1/2)/N}]}{e^{-jk(2\pi/2N)}[e^{jk(2\pi/2N)} - e^{-jk(2\pi/2N)}]} \\ &= \frac{1}{N} \frac{\sin[2\pi k(N_1 + 1/2)/N]}{\sin(\pi k/N)}, \quad k \neq 0, \pm N, \pm 2N, \ldots \end{aligned} \tag{3.104}$$

and

$$a_k = \frac{2N_1 + 1}{N}, \quad k = 0, \pm N, \pm 2N, \ldots. \tag{3.105}$$

The coefficients a_k for $2N_1 + 1 = 5$ are sketched for $N = 10$, 20, and 40 in Figures 3.17(a), (b), and (c), respectively.

In discussing the convergence of the continuous-time Fourier series in Section 3.4, we considered the example of a symmetric square wave and observed how the finite sum in eq. (3.52) converged to the square wave as the number of terms approached infinity. In par-

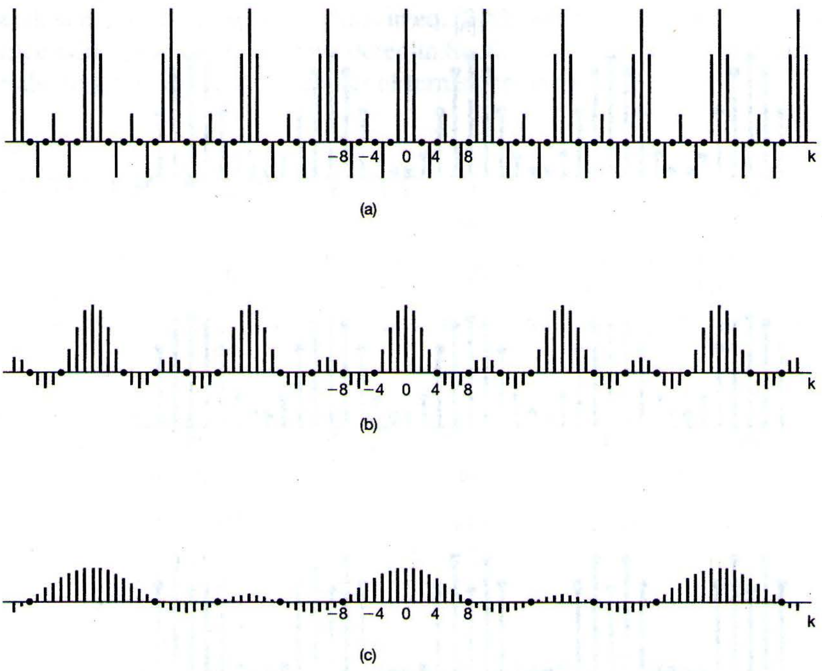


Figure 3.17 Fourier series coefficients for the periodic square wave of Example 3.12; plots of Na_k for $2N_1 + 1 = 5$ and (a) $N = 10$; (b) $N = 20$; and (c) $N = 40$.

ticular, we observed the Gibbs phenomenon at the discontinuity, whereby, as the number of terms increased, the ripples in the partial sum (Figure 3.9) became compressed toward the discontinuity, with the peak amplitude of the ripples remaining constant independently of the number of terms in the partial sum. Let us consider the analogous sequence of partial sums for the discrete-time square wave, where, for convenience, we will assume that the period N is odd. In Figure 3.18, we have depicted the signals

$$\hat{x}[n] = \sum_{k=-M}^{M} a_k e^{jk(2\pi/N)n} \tag{3.106}$$

for the example of Figure 3.16 with $N = 9$, $2N_1 + 1 = 5$, and for several values of M. For $M = 4$, the partial sum exactly equals $x[n]$. We see in particular that in contrast to the continuous-time case, there are no convergence issues and there is no Gibbs phenomenon. In fact, there are no convergence issues with the discrete-time Fourier series in general. The reason for this stems from the fact that any discrete-time periodic sequence $x[n]$ is completely specified by a *finite* number N of parameters, namely, the values of the sequence over one period. The Fourier series analysis equation (3.95) simply transforms this set of N parameters into an equivalent set—the values of the N Fourier coefficients—and

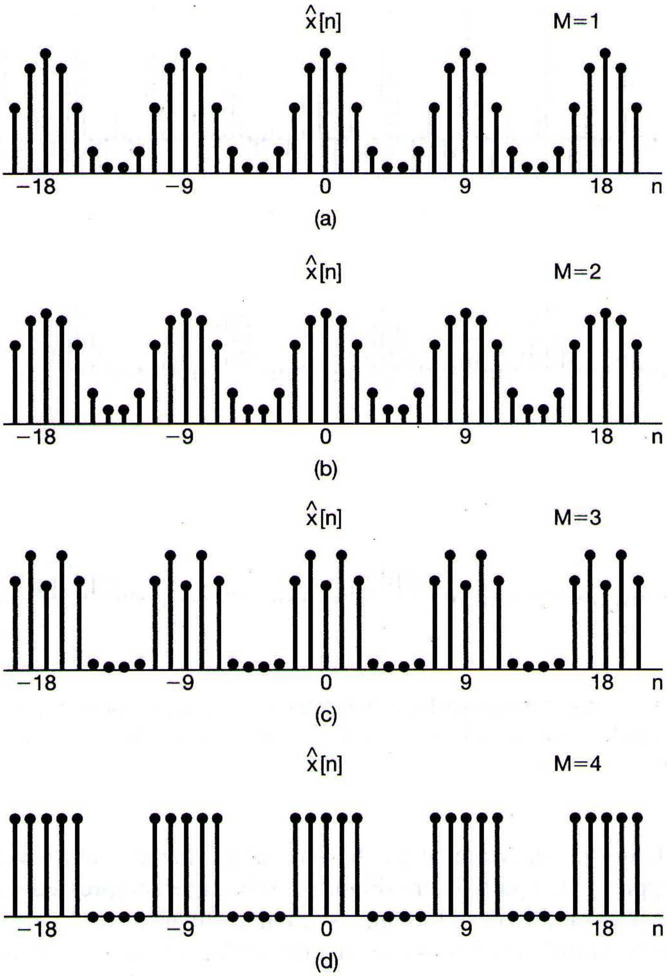


Figure 3.18 Partial sums of eqs. (3.106) and (3.107) for the periodic square wave of Figure 3.16 with $N = 9$ and $2N_1 + 1 = 5$: (a) $M = 1$; (b) $M = 2$; (c) $M = 3$; (d) $M = 4$.

the synthesis equation (3.94) tells us how to recover the values of the original sequence in terms of a *finite* series. Thus, if N is odd and we take $M = (N - 1)/2$ in eq. (3.106), the sum includes exactly N terms, and consequently, from the synthesis equations, we have $\hat{x}[n] = x[n]$. Similarly, if N is even and we let

$$\hat{x}[n] = \sum_{k=-M+1}^{M} a_k e^{jk(2\pi/N)n}, \tag{3.107}$$

then with $M = N/2$, this sum consists of N terms, and again, we can conclude from eq. (3.94) that $\hat{x}[n] = x[n]$.

In contrast, a continuous-time periodic signal takes on a continuum of values over a single period, and an infinite number of Fourier coefficients are required to represent it.

Thus, in general, *none* of the finite partial sums in eq. (3.52) yield the exact values of $x(t)$, and convergence issues, such as those considered in Section 3.4, arise as we consider the problem of evaluating the limit as the number of terms approaches infinity.

3.7 PROPERTIES OF DISCRETE-TIME FOURIER SERIES

There are strong similarities between the properties of discrete-time and continuous-time Fourier series. This can be readily seen by comparing the discrete-time Fourier series properties summarized in Table 3.2 with their continuous-time counterparts in Table 3.1.

TABLE 3.2 PROPERTIES OF DISCRETE-TIME FOURIER SERIES

Property	**Periodic Signal**	**Fourier Series Coefficients**
	$\left.\begin{matrix} x[n] \\ y[n] \end{matrix}\right\}$ Periodic with period N and fundamental frequency $\omega_0 = 2\pi/N$	$\left.\begin{matrix} a_k \\ b_k \end{matrix}\right\}$ Periodic with period N
Linearity	$Ax[n] + By[n]$	$Aa_k + Bb_k$
Time Shifting	$x[n - n_0]$	$a_k e^{-jk(2\pi/N)n_0}$
Frequency Shifting	$e^{jM(2\pi/N)n} x[n]$	a_{k-M}
Conjugation	$x^*[n]$	a^*_{-k}
Time Reversal	$x[-n]$	a_{-k}
Time Scaling	$x_{(m)}[n] = \begin{cases} x[n/m], & \text{if } n \text{ is a multiple of } m \\ 0, & \text{if } n \text{ is not a multiple of } m \end{cases}$ (periodic with period mN)	$\frac{1}{m} a_k \begin{pmatrix} \text{viewed as periodic} \\ \text{with period } mN \end{pmatrix}$
Periodic Convolution	$\sum_{r=\langle N \rangle} x[r]y[n-r]$	$Na_k b_k$
Multiplication	$x[n]y[n]$	$\sum_{l=\langle N \rangle} a_l b_{k-l}$
First Difference	$x[n] - x[n-1]$	$(1 - e^{-jk(2\pi/N)})a_k$
Running Sum	$\sum_{k=-\infty}^{n} x[k] \begin{pmatrix} \text{finite valued and periodic only} \\ \text{if } a_0 = 0 \end{pmatrix}$	$\left(\frac{1}{(1 - e^{-jk(2\pi/N)})} \right) a_k$
Conjugate Symmetry for Real Signals	$x[n]$ real	$\begin{cases} a_k = a^*_{-k} \\ \Re e\{a_k\} = \Re e\{a_{-k}\} \\ \Im m\{a_k\} = -\Im m\{a_{-k}\} \\ \lvert a_k \rvert = \lvert a_{-k} \rvert \\ \sphericalangle a_k = -\sphericalangle a_{-k} \end{cases}$
Real and Even Signals	$x[n]$ real and even	a_k real and even
Real and Odd Signals	$x[n]$ real and odd	a_k purely imaginary and odd
Even-Odd Decomposition of Real Signals	$\begin{cases} x_e[n] = \mathcal{E}v\{x[n]\} & [x[n] \text{ real}] \\ x_o[n] = \mathcal{O}d\{x[n]\} & [x[n] \text{ real}] \end{cases}$	$\Re e\{a_k\}$ $j\Im m\{a_k\}$

Parseval's Relation for Periodic Signals

$$\frac{1}{N} \sum_{n=\langle N \rangle} |x[n]|^2 = \sum_{k=\langle N \rangle} |a_k|^2$$

The derivations of many of these properties are very similar to those of the corresponding properties for continuous-time Fourier series, and several such derivations are considered in the problems at the end of the chapter. In addition, in Chapter 5 we will see that most of the properties can be inferred from corresponding properties of the discrete-time Fourier transform. Consequently, we limit the discussion in the following subsections to only a few of these properties, including several that have important differences relative to those for continuous time. We also provide examples illustrating the usefulness of various discrete-time Fourier series properties for developing conceptual insights and helping to reduce the complexity of the evaluation of the Fourier series of many periodic sequences.

As with continuous time, it is often convenient to use a shorthand notation to indicate the relationship between a periodic signal and its Fourier series coefficients. Specifically, if $x[n]$ is a periodic signal with period N and with Fourier series coefficients denoted by a_k, then we will write

$$x[n] \overset{\mathcal{FS}}{\longleftrightarrow} a_k.$$

3.7.1 Multiplication

The multiplication property of the Fourier series representation is one example of a property that reflects the difference between continuous time and discrete time. From Table 3.1, the product of two continuous-time signals of period T results in a periodic signal with period T whose sequence of Fourier series coefficients is the *convolution* of the sequences of Fourier series coefficients of the two signals being multiplied. In discrete time, suppose that

$$x[n] \overset{\mathcal{FS}}{\longleftrightarrow} a_k$$

and

$$y[n] \overset{\mathcal{FS}}{\longleftrightarrow} b_k$$

are both periodic with period N. Then the product $x[n]y[n]$ is also periodic with period N, and, as shown in Problem 3.57, its Fourier coefficients, d_k, are given by

$$x[n]y[n] \overset{\mathcal{FS}}{\longleftrightarrow} d_k = \sum_{l=\langle N\rangle} a_l b_{k-l}. \tag{3.108}$$

Equation (3.108) is analogous to the definition of convolution, except that the summation variable is now restricted to an interval of N consecutive samples. As shown in Problem 3.57, the summation can be taken over *any* set of N consecutive values of l. We refer to this type of operation as a *periodic convolution* between the two periodic sequences of Fourier coefficients. The usual form of the convolution sum (where the summation variable ranges from $-\infty$ to ∞) is sometimes referred to as *aperiodic convolution,* to distinguish it from periodic convolution.

3.7.2 First Difference

The discrete-time parallel to the differentiation property of the continuous-time Fourier series involves the use of the first-difference operation, which is defined as $x[n]-x[n-1]$.

If $x[n]$ is periodic with period N, then so is $y[n]$, since shifting $x[n]$ or linearly combining $x[n]$ with another periodic signal whose period is N always results in a periodic signal with period N. Also, if

$$x[n] \overset{\mathcal{FS}}{\longleftrightarrow} a_k,$$

then the Fourier coefficients corresponding to the first difference of $x[n]$ may be expressed as

$$x[n] - x[n-1] \overset{\mathcal{FS}}{\longleftrightarrow} (1 - e^{-jk(2\pi/N)})a_k, \tag{3.109}$$

which is easily obtained by applying the time-shifting and linearity properties in Table 3.2. A common use of this property is in situations where evaluation of the Fourier series coefficients is easier for the first difference than for the original sequence. (See Problem 3.31.)

3.7.3 Parseval's Relation for Discrete-Time Periodic Signals

As shown in Problem 3.57, Parseval's relation for discrete-time periodic signals is given by

$$\frac{1}{N}\sum_{n=\langle N\rangle} |x[n]|^2 = \sum_{k=\langle N\rangle} |a_k|^2, \tag{3.110}$$

where the a_k are the Fourier series coefficients of $x[n]$ and N is the period. As in the continuous-time case, the left-hand side of Parseval's relation is the average power in one period for the periodic signal $x[n]$. Similarly, $|a_k|^2$ is the average power in the kth harmonic component of $x[n]$. Thus, once again, Parseval's relation states that the average power in a periodic signal equals the sum of the average powers in all of its harmonic components. In discrete time, of course, there are only N distinct harmonic components, and since the a_k are periodic with period N, the sum on the right-hand side of eq. (3.110) can be taken over any N consecutive values of k.

3.7.4 Examples

In this subsection, we present several examples illustrating how properties of the discrete-time Fourier series can be used to characterize discrete-time periodic signals and to compute their Fourier series representations. Specifically, Fourier series properties, such as those listed in Table 3.2, may be used to simplify the process of determining the Fourier series coefficients of a given signal. This involves first expressing the given signal in terms of other signals whose Fourier series coefficients are already known or are simpler to compute. Then, using Table 3.2, we can express the Fourier series coefficients of the given signal in terms of the Fourier series coefficients of the other signals. This is illustrated in Example 3.13. Example 3.14 then illustrates the determination of a sequence from some partial information. In Example 3.15 we illustrate the use of the periodic convolution property in Table 3.2.

Example 3.13

Let us consider the problem of finding the Fourier series coefficients a_k of the sequence $x[n]$ shown in Figure 3.19(a). This sequence has a fundamental period of 5. We observe that $x[n]$ may be viewed as the sum of the square wave $x_1[n]$ in Figure 3.19(b) and the *dc sequence* $x_2[n]$ in Figure 3.19(c). Denoting the Fourier series coefficients of $x_1[n]$ by b_k and those of $x_2[n]$ by c_k, we use the linearity property of Table 3.2 to conclude that

$$a_k = b_k + c_k. \tag{3.111}$$

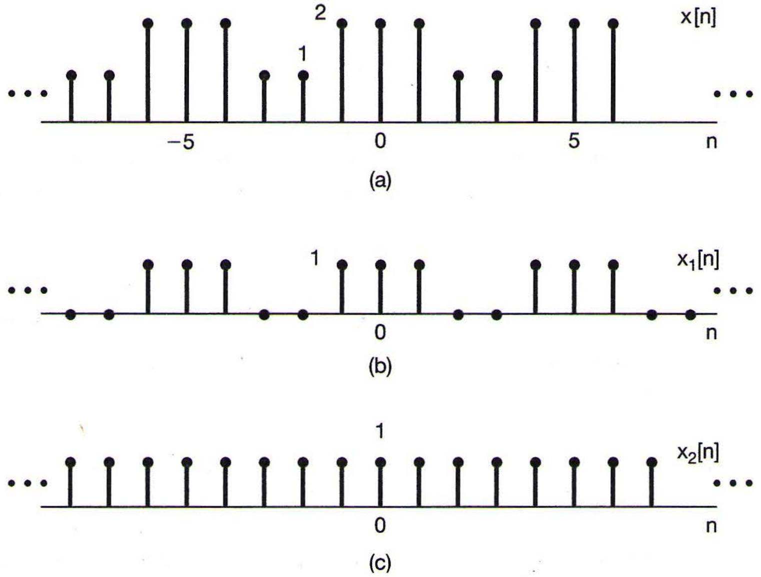


Figure 3.19 (a) Periodic sequence $x[n]$ for Example 3.13 and its representation as a sum of (b) the square wave $x_1[n]$ and (c) the dc sequence $x_2[n]$.

From Example 3.12 (with $N_1 = 1$ and $N = 5$), the Fourier series coefficients b_k corresponding to $x_1[n]$ can be expressed as

$$b_k = \begin{cases} \dfrac{1}{5}\dfrac{\sin(3\pi k/5)}{\sin(\pi k/5)}, & \text{for } k \neq 0, \pm 5, \pm 10, \ldots \\ \dfrac{3}{5}, & \text{for } k = 0, \pm 5, \pm 10, \ldots \end{cases}. \tag{3.112}$$

The sequence $x_2[n]$ has only a dc value, which is captured by its zeroth Fourier series coefficient:

$$c_0 = \frac{1}{5}\sum_{n=0}^{4} x_2[n] = 1. \tag{3.113}$$

Since the discrete-time Fourier series coefficients are periodic, it follows that $c_k = 1$ whenever k is an integer multiple of 5. The remaining coefficients of $x_2[n]$ must be zero, because $x_2[n]$ contains only a dc component. We can now substitute the expressions for b_k and c_k into eq. (3.111) to obtain

$$a_k = \begin{cases} b_k = \dfrac{1}{5}\dfrac{\sin(3\pi k/5)}{\sin(\pi k/5)}, & \text{for } k \neq 0, \pm 5, \pm 10, \ldots \\ \dfrac{8}{5}, & \text{for } k = 0, \pm 5, \pm 10, \ldots \end{cases}. \tag{3.114}$$

Example 3.14

Suppose we are given the following facts about a sequence $x[n]$:

1. $x[n]$ is periodic with period $N = 6$.
2. $\sum_{n=0}^{5} x[n] = 2$.
3. $\sum_{n=2}^{7} (-1)^n x[n] = 1$.
4. $x[n]$ has the minimum power per period among the set of signals satisfying the preceding three conditions.

Let us determine the sequence $x[n]$. We denote the Fourier series coefficients of $x[n]$ by a_k. From Fact 2, we conclude that $a_0 = 1/3$. Noting that $(-1)^n = e^{-j\pi n} = e^{-j(2\pi/6)3n}$, we see from Fact 3 that $a_3 = 1/6$. From Parseval's relation (see Table 3.2), the average power in $x[n]$ is

$$P = \sum_{k=0}^{5} |a_k|^2. \tag{3.115}$$

Since each nonzero coefficient contributes a positive amount to P, and since the values of a_0 and a_3 are prespecified, the value of P is minimized by choosing $a_1 = a_2 = a_4 = a_5 = 0$. It then follows that

$$x[n] = a_0 + a_3 e^{j\pi n} = (1/3) + (1/6)(-1)^n, \tag{3.116}$$

which is sketched in Figure 3.20.

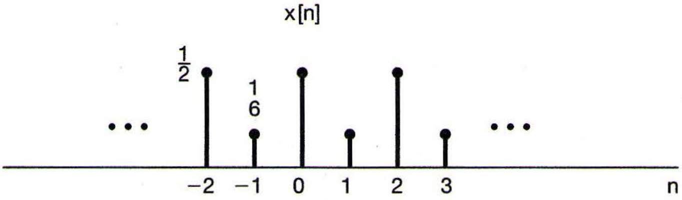


Figure 3.20 Sequence $x[n]$ that is consistent with the properties specified in Example 3.14.

Example 3.15

In this example we determine and sketch a periodic sequence, given an algebraic expression for its Fourier series coefficients. In the process, we will also exploit the periodic convolution property (see Table 3.2) of the discrete-time Fourier series. Specifically, as stated in the table and as shown in Problem 3.58, if $x[n]$ and $y[n]$ are periodic with period N, then the signal

$$w[n] = \sum_{r=\langle N \rangle} x[r]y[n-r]$$

is also periodic with period N. Here, the summation may be taken over any set of N consecutive values of r. Furthermore, the Fourier series coefficients of $w[n]$ are equal to Na_kb_k, where a_k and b_k are the Fourier coefficients of $x[n]$ and $y[n]$, respectively.

Suppose now that we are told that a signal $w[n]$ is periodic with a fundamental period of $N = 7$ and with Fourier series coefficients

$$c_k = \frac{\sin^2(3\pi k/7)}{7\sin^2(\pi k/7)}. \tag{3.117}$$

We observe that $c_k = 7d_k^2$, where d_k denotes the sequence of Fourier series coefficients of a square wave $x[n]$, as in Example 3.12, with $N_1 = 1$ and $N = 7$. Using the periodic convolution property, we see that

$$w[n] = \sum_{r=\langle 7\rangle} x[r]x[n-r] = \sum_{r=-3}^{3} x[r]x[n-r], \tag{3.118}$$

where, in the last equality, we have chosen to sum over the interval $-3 \le r \le 3$. Except for the fact that the sum is limited to a finite interval, the product-and-sum method for evaluating convolution is applicable here. In fact, we can convert eq. (3.118) to an ordinary convolution by defining a signal $\hat{x}[n]$ that equals $x[n]$ for $-3 \le n \le 3$ and is zero otherwise. Then, from eq. (3.118),

$$w[n] = \sum_{r=-3}^{3} \hat{x}[r]x[n-r] = \sum_{r=-\infty}^{+\infty} \hat{x}[r]x[n-r].$$

That is, $w[n]$ is the aperiodic convolution of the sequences $\hat{x}[n]$ and $x[n]$.

The sequences $x[r]$, $\hat{x}[r]$, and $x[n-r]$ are sketched in Figure 3.21 (a)–(c). From the figure we can immediately calculate $w[n]$. In particular we see that $w[0] = 3$; $w[-1] = w[1] = 2$; $w[-2] = w[2] = 1$; and $w[-3] = w[3] = 0$. Since $w[n]$ is periodic with period 7, we can then sketch $w[n]$ as shown in Figure 3.21(d).

3.8 FOURIER SERIES AND LTI SYSTEMS

In the preceding few sections, we have seen that the Fourier series representation can be used to construct any periodic signal in discrete time and essentially all periodic continuous-time signals of practical importance. In addition, in Section 3.2 we saw that the response of an LTI system to a linear combination of complex exponentials takes a particularly simple form. Specifically, in continuous time, if $x(t) = e^{st}$ is the input to a continuous-time LTI system, then the output is given by $y(t) = H(s)e^{st}$, where, from eq. (3.6),

$$H(s) = \int_{-\infty}^{+\infty} h(\tau)e^{-s\tau}d\tau, \tag{3.119}$$

in which $h(t)$ is the impulse response of the LTI system.

Similarly, if $x[n] = z^n$ is the input to a discrete-time LTI system, then the output is given by $y[n] = H(z)z^n$, where, from eq. (3.10),

$$H(z) = \sum_{k=-\infty}^{+\infty} h[k]z^{-k}, \tag{3.120}$$

in which $h[n]$ is the impulse response of the LTI system.

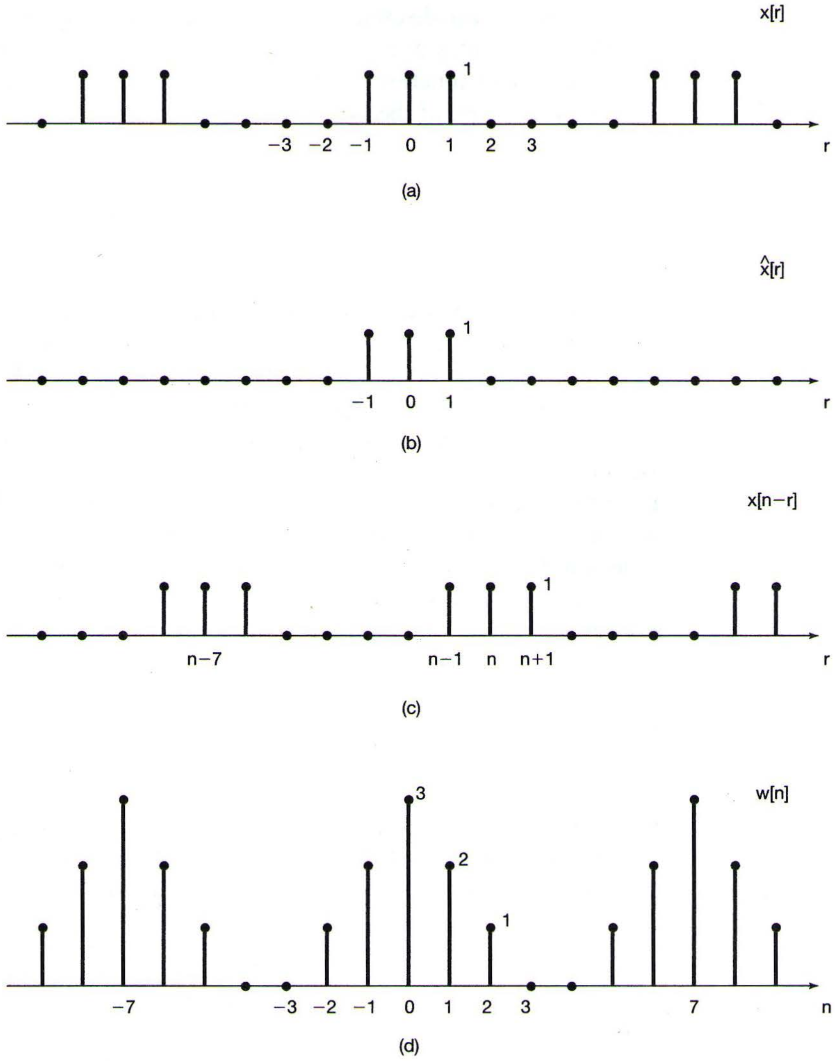


Figure 3.21 (a) The square-wave sequence $x[r]$ in Example 3.15; (b) the sequence $\hat{x}[r]$ equal to $x[r]$ for $-3 \leq r \leq 3$ and zero otherwise; (c) the sequence $x[n-r]$; (d) the sequence $w[n]$ equal to the periodic convolution of $x[n]$ with itself and to the aperiodic convolution of $\hat{x}[n]$ with $x[n]$.

When s or z are general complex numbers, $H(s)$ and $H(z)$ are referred to as the *system functions* of the corresponding systems. For continuous-time signals and systems in this and the following chapter, we focus on the specific case in which $\mathcal{R}e\{s\} = 0$, so that $s = j\omega$, and consequently, e^{st} is of the form $e^{j\omega t}$. This input is a complex exponential at frequency ω. The system function of the form $s = j\omega$—i.e., $H(j\omega)$ viewed as a function of ω—is referred to as the *frequency response* of the system and is given by

$$H(j\omega) = \int_{-\infty}^{+\infty} h(t)e^{-j\omega t}dt. \tag{3.121}$$

Similarly, for discrete-time signals and systems, we focus in this chapter and in Chapter 5 on values of z for which $|z| = 1$, so that $z = e^{j\omega}$ and z^n is of the form $e^{j\omega n}$. Then the system function $H(z)$ for z restricted to the form $z = e^{j\omega}$ is referred to as the frequency response of the system and is given by

$$H(e^{j\omega}) = \sum_{n=-\infty}^{+\infty} h[n]e^{-j\omega n}. \tag{3.122}$$

The response of an LTI system to a complex exponential signal of the form $e^{j\omega t}$ (in continuous time) or $e^{j\omega n}$ (in discrete time) is particularly simple to express in terms of the frequency response of the system. Furthermore, as a result of the superposition property for LTI systems, we can express the response of an LTI system to a linear combination of complex exponentials with equal ease. In Chapters 4 and 5, we will see how we can use these ideas together with continuous-time and discrete-time Fourier transforms to analyze the response of LTI systems to aperiodic signals. In the remainder of this chapter, as a first look at this important set of concepts and results, we focus on interpreting and understanding this notion in the context of periodic signals.

Consider first the continuous-time case, and let $x(t)$ be a periodic signal with a Fourier series representation given by

$$x(t) = \sum_{k=-\infty}^{+\infty} a_k e^{jk\omega_0 t}. \tag{3.123}$$

Suppose that we apply this signal as the input to an LTI system with impulse response $h(t)$. Then, since each of the complex exponentials in eq. (3.123) is an eigenfunction of the system, as in eq. (3.13) with $s_k = jk\omega_0$, it follows that the output is

$$y(t) = \sum_{k=-\infty}^{+\infty} a_k H(jk\omega_0) e^{jk\omega_0 t}. \tag{3.124}$$

Thus, $y(t)$ is also periodic with the same fundamental frequency as $x(t)$. Furthermore, if $\{a_k\}$ is the set of Fourier series coefficients for the input $x(t)$, then $\{a_k H(jk\omega_0)\}$ is the set of coefficients for the output $y(t)$. That is, the effect of the LTI system is to modify individually each of the Fourier coefficients of the input through multiplication by the value of the frequency response at the corresponding frequency.

Example 3.16

Suppose that the periodic signal $x(t)$ discussed in Example 3.2 is the input signal to an LTI system with impulse response

$$h(t) = e^{-t}u(t).$$

To calculate the Fourier series coefficients of the output $y(t)$, we first compute the frequency response:

$$\begin{aligned} H(j\omega) &= \int_0^{\infty} e^{-\tau} e^{-j\omega\tau}\, d\tau \\ &= -\frac{1}{1+j\omega} e^{-\tau} e^{-j\omega\tau}\Big|_0^{\infty} \\ &= \frac{1}{1+j\omega}. \end{aligned} \tag{3.125}$$

Therefore, using eqs. (3.124) and (3.125), together with the fact that $\omega_0 = 2\pi$ in this example, we obtain

$$y(t) = \sum_{k=-3}^{+3} b_k e^{jk2\pi t}, \tag{3.126}$$

with $b_k = a_k H(jk2\pi)$, so that

$$\begin{aligned} &b_0 = 1, \\ b_1 &= \frac{1}{4}\left(\frac{1}{1+j2\pi}\right), \quad & b_{-1} &= \frac{1}{4}\left(\frac{1}{1-j2\pi}\right), \\ b_2 &= \frac{1}{2}\left(\frac{1}{1+j4\pi}\right), \quad & b_{-2} &= \frac{1}{2}\left(\frac{1}{1-j4\pi}\right), \\ b_3 &= \frac{1}{3}\left(\frac{1}{1+j6\pi}\right), \quad & b_{-3} &= \frac{1}{3}\left(\frac{1}{1-j6\pi}\right). \end{aligned} \tag{3.127}$$

Note that $y(t)$ must be a real-valued signal, since it is the convolution of $x(t)$ and $h(t)$, which are both real. This can be verified by examining eq. (3.127) and observing that $b_k^* = b_{-k}$. Therefore, $y(t)$ can also be expressed in either of the forms given in eqs. (3.31) and (3.32); that is,

$$y(t) = 1 + 2\sum_{k=1}^{3} D_k \cos(2\pi kt + \theta_k), \tag{3.128}$$

or

$$y(t) = 1 + 2\sum_{k=1}^{3} [E_k \cos 2\pi kt - F_k \sin 2\pi kt], \tag{3.129}$$

where

$$b_k = D_k e^{j\theta_k} = E_k + jF_k, \quad k = 1, 2, 3. \tag{3.130}$$

These coefficients can be evaluated directly from eq. (3.127). For example,

$$\begin{aligned} D_1 &= |b_1| = \frac{1}{4\sqrt{1+4\pi^2}}, \quad & \theta_1 &= \sphericalangle b_1 = -\tan^{-1}(2\pi), \\ E_1 &= \mathcal{R}e\{b_1\} = \frac{1}{4(1+4\pi^2)}, \quad & F_1 &= \mathcal{I}m\{b_1\} = -\frac{\pi}{2(1+4\pi^2)}. \end{aligned}$$

In discrete time, the relationship between the Fourier series coefficients of the input and output of an LTI system exactly parallels eqs. (3.123) and (3.124). Specifically, let $x[n]$ be a periodic signal with Fourier series representation given by

$$x[n] = \sum_{k=\langle N\rangle} a_k e^{jk(2\pi/N)n}.$$

If we apply this signal as the input to an LTI system with impulse response $h[n]$, then, as in eq. (3.16) with $z_k = e^{jk(2\pi/N)}$, the output is

$$y[n] = \sum_{k=\langle N\rangle} a_k H(e^{j2\pi k/N}) e^{jk(2\pi/N)n}. \tag{3.131}$$

Thus, $y[n]$ is also periodic with the same period as $x[n]$, and the kth Fourier coefficient of $y[n]$ is the product of the kth Fourier coefficient of the input and the value of the frequency response of the LTI system, $H(e^{j2\pi k/N})$, at the corresponding frequency.

Example 3.17

Consider an LTI system with impulse response $h[n] = \alpha^n u[n]$, $-1 < \alpha < 1$, and with the input

$$x[n] = \cos\left(\frac{2\pi n}{N}\right). \tag{3.132}$$

As in Example 3.10, $x[n]$ can be written in Fourier series form as

$$x[n] = \frac{1}{2}e^{j(2\pi/N)n} + \frac{1}{2}e^{-j(2\pi/N)n}.$$

Also, from eq. (3.122),

$$H(e^{j\omega}) = \sum_{n=0}^{\infty} \alpha^n e^{-j\omega n} = \sum_{n=0}^{\infty} \left(\alpha e^{-j\omega}\right)^n. \tag{3.133}$$

This geometric series can be evaluated using the result of Problem 1.54, yielding

$$H(e^{j\omega}) = \frac{1}{1 - \alpha e^{-j\omega}}. \tag{3.134}$$

Using eq. (3.131), we then obtain the Fourier series for the output:

$$\begin{aligned} y[n] &= \frac{1}{2}H\left(e^{j2\pi/N}\right)e^{j(2\pi/N)n} + \frac{1}{2}H\left(e^{-j2\pi/N}\right)e^{-j(2\pi/N)n} \\ &= \frac{1}{2}\left(\frac{1}{1 - \alpha e^{-j2\pi/N}}\right)e^{j(2\pi/N)n} + \frac{1}{2}\left(\frac{1}{1 - \alpha e^{j2\pi/N}}\right)e^{-j(2\pi/N)n}. \end{aligned} \tag{3.135}$$

If we write

$$\frac{1}{1 - \alpha e^{-j2\pi/N}} = re^{j\theta},$$

then eq. (3.135) reduces to

$$y[n] = r\cos\left(\frac{2\pi}{N}n + \theta\right). \tag{3.136}$$

For example, if $N = 4$,

$$\frac{1}{1 - \alpha e^{-j2\pi/4}} = \frac{1}{1 + \alpha j} = \frac{1}{\sqrt{1 + \alpha^2}} e^{j(-\tan^{-1}(\alpha))},$$

and thus,

$$y[n] = \frac{1}{\sqrt{1 + \alpha^2}} \cos\left(\frac{\pi n}{2} - \tan^{-1}(\alpha)\right).$$

We note that for expressions such as eqs. (3.124) and (3.131) to make sense, the frequency responses $H(j\omega)$ and $H(e^{j\omega})$ in eqs. (3.121) and (3.122) must be well defined and finite. As we will see in Chapters 4 and 5, this will be the case if the LTI systems under consideration are stable. For example, the LTI system in Example 3.16, with impulse response $h(t) = e^{-t}u(t)$, is stable and has a well-defined frequency response given by eq. (3.125). On the other hand, an LTI system with impulse response $h(t) = e^{t}u(t)$ is unstable, and it is easy to check that the integral in eq. (3.121) for $H(j\omega)$ diverges for any value of ω. Similarly, the LTI system in Example 3.17, with impulse response $h[n] = \alpha^n u[n]$, is stable for $|\alpha| < 1$ and has frequency response given by eq. (3.134). However, if $|\alpha| > 1$, the system is unstable, and then the summation in eq. (3.133) diverges.

3.9 FILTERING

In a variety of applications, it is of interest to change the relative amplitudes of the frequency components in a signal or perhaps eliminate some frequency components entirely, a process referred to as *filtering*. Linear time-invariant systems that change the shape of the spectrum are often referred to as *frequency-shaping filters*. Systems that are designed to pass some frequencies essentially undistorted and significantly attenuate or eliminate others are referred to as *frequency-selective filters*. As indicated by eqs. (3.124) and (3.131), the Fourier series coefficients of the output of an LTI system are those of the input multiplied by the frequency response of the system. Consequently, filtering can be conveniently accomplished through the use of LTI systems with an appropriately chosen frequency response, and frequency-domain methods provide us with the ideal tools to examine this very important class of applications. In this and the following two sections, we take a first look at filtering through a few examples.

3.9.1 Frequency-Shaping Filters

One application in which frequency-shaping filters are often encountered is audio systems. For example, LTI filters are typically included in such systems to permit the listener to modify the relative amounts of low-frequency energy (bass) and high-frequency energy (treble). These filters correspond to LTI systems whose frequency responses can be changed by manipulating the tone controls. Also, in high-fidelity audio systems, a so-called equalizing filter is often included in the preamplifier to compensate for the frequency-response characteristics of the speakers. Overall, these cascaded filtering stages are frequently referred to as the equalizing or equalizer circuits for the audio system. Figure 3.22 illustrates the three stages of the equalizer circuits for one particular series of audio speakers. In this figure, the magnitude of the frequency response for each of these stages is shown on a log-log plot. Specifically, the magnitude is in units of $20\log_{10}|H(j\omega)|$, referred to as decibels or dB. The frequency axis is labeled in Hz (i.e., $\omega/2\pi$) along a logarithmic scale. As will be discussed in more detail in Section 6.2.3, a logarithmic display of the magnitude of the frequency response in this form is common and useful.

Taken together, the equalizing circuits in Figure 3.22 are designed to compensate for the frequency response of the speakers and the room in which they are located and to allow the listener to control the overall frequency response. In particular, since the three systems are connected in cascade, and since each system modifies a complex exponential input $Ke^{j\omega t}$ by multiplying it by the system frequency response at that frequency, it follows that the overall frequency response of the cascade of the three systems is the product of the three frequency responses. The first two filters, indicated in Figures 3.22(a) and (b), together make up the control stage of the system, as the frequency behavior of these filters can be adjusted by the listener. The third filter, illustrated in Figure 3.22(c), is the equalizer stage, which has the fixed frequency response indicated. The filter in Figure 3.22(a) is a low-frequency filter controlled by a two-position switch, to provide one of the two frequency responses indicated. The second filter in the control stage has two continuously adjustable slider switches to vary the frequency response within the limits indicated in Figure 3.22(b).

Another class of frequency-shaping filters often encountered is that for which the filter output is the derivative of the filter input, i.e., $y(t) = dx(t)/dt$. With $x(t)$ of the form $x(t) = e^{j\omega t}$, $y(t)$ will be $y(t) = j\omega e^{j\omega t}$, from which it follows that the frequency response is

$$H(j\omega) = j\omega. \tag{3.137}$$

The frequency response characteristics of a differentiating filter are shown in Figure 3.23. Since $H(j\omega)$ is complex in general, and in this example in particular, $H(j\omega)$ is frequently displayed (as in the figure) as separate plots of $|H(j\omega)|$ and $\sphericalangle H(j\omega)$. The shape of this frequency response implies that a complex exponential input $e^{j\omega t}$ will receive greater amplification for larger values of ω. Consequently, differentiating filters are useful in enhancing rapid variations or transitions in a signal.

One purpose for which differentiating filters are often used is to enhance edges in picture processing. A black-and-white picture can be thought of as a two-dimensional "continuous-time" signal $x(t_1, t_2)$, where t_1 and t_2 are the horizontal and vertical coordinates, respectively, and $x(t_1, t_2)$ is the brightness of the image. If the image is repeated periodically in the horizontal and vertical directions, then it can be represented by a two-dimensional Fourier series (see Problem 3.70) consisting of sums of products of complex

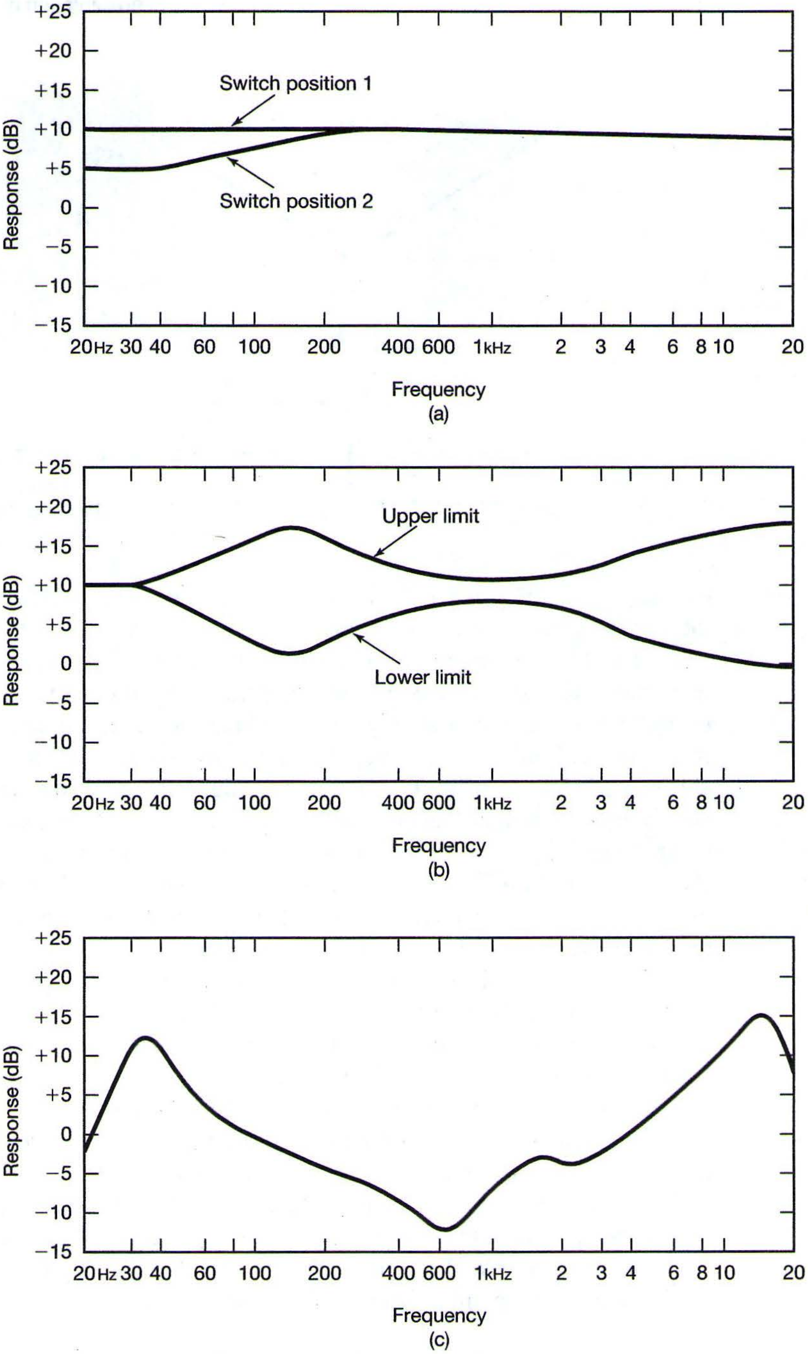


Figure 3.22 Magnitudes of the frequency responses of the equalizer circuits for one particular series of audio speakers, shown on a scale of $20 \log_{10} |H(j\omega)|$, which is referred to as a decibel (or dB) scale. (a) Low-frequency filter controlled by a two-position switch; (b) upper and lower frequency limits on a continuously adjustable shaping filter; (c) fixed frequency response of the equalizer stage.

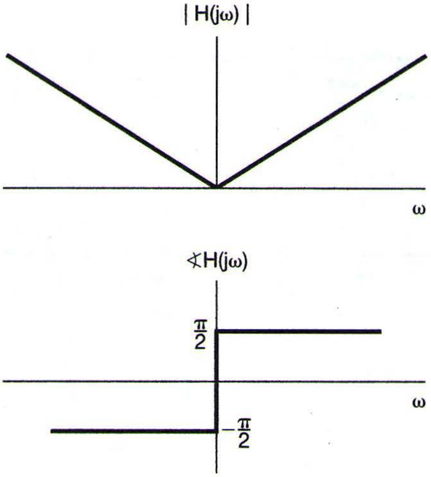


Figure 3.23 Characteristics of the frequency response of a filter for which the output is the derivative of the input.

exponentials, $e^{j\omega_1 t_1}$ and $e^{j\omega_2 t_2}$, that oscillate at possibly different frequencies in each of the two coordinate directions. Slow variations in brightness in a particular direction are represented by the lower harmonics in that direction. For example, consider an edge corresponding to a sharp transition in brightness that runs vertically in an image. Since the brightness is constant or slowly varying along the edge, the frequency content of the edge in the vertical direction is concentrated at low frequencies. In contrast, since there is an abrupt variation in brightness across the edge, the frequency content of the edge in the horizontal direction is concentrated at higher frequencies. Figure 3.24 illustrates the effect on an image of the two-dimensional equivalent of a differentiating filter.[11] Figure 3.24(a) shows two original images and Figure 3.24(b) the result of processing those images with the filter. Since the derivative at the edges of a picture is greater than in regions where the brightness varies slowly with distance, the effect of the filter is to enhance the edges.

Discrete-time LTI filters also find a broad array of applications. Many of these involve the use of discrete-time systems, implemented using general- or special-purpose digital processors, to process continuous-time signals, a topic we discuss at some length in Chapter 7. In addition, the analysis of time series information, including demographic data and economic data sequences such as the stock market average, commonly involves the use of discrete-time filters. Often the long-term variations (which correspond to low frequencies) have a different significance than the short-term variations (which correspond to high frequencies), and it is useful to analyze these components separately. Reshaping the relative weighting of the components is typically accomplished using discrete-time filters.

As one example of a simple discrete-time filter, consider an LTI system that successively takes a two-point average of the input values:

$$y[n] = \frac{1}{2}(x[n] + x[n-1]). \tag{3.138}$$

---

[11] Specifically each image in Figure 3.24(b) is the magnitude of the two-dimensional gradient of its counterpart image in Figure 3.24(a) where the magnitude of the gradient of $f(x, y)$ is

$$\left[\left(\frac{\partial f(x,y)}{\partial x}\right)^2 + \left(\frac{\partial f(x,y)}{\partial y}\right)^2\right]^{1/2}$$

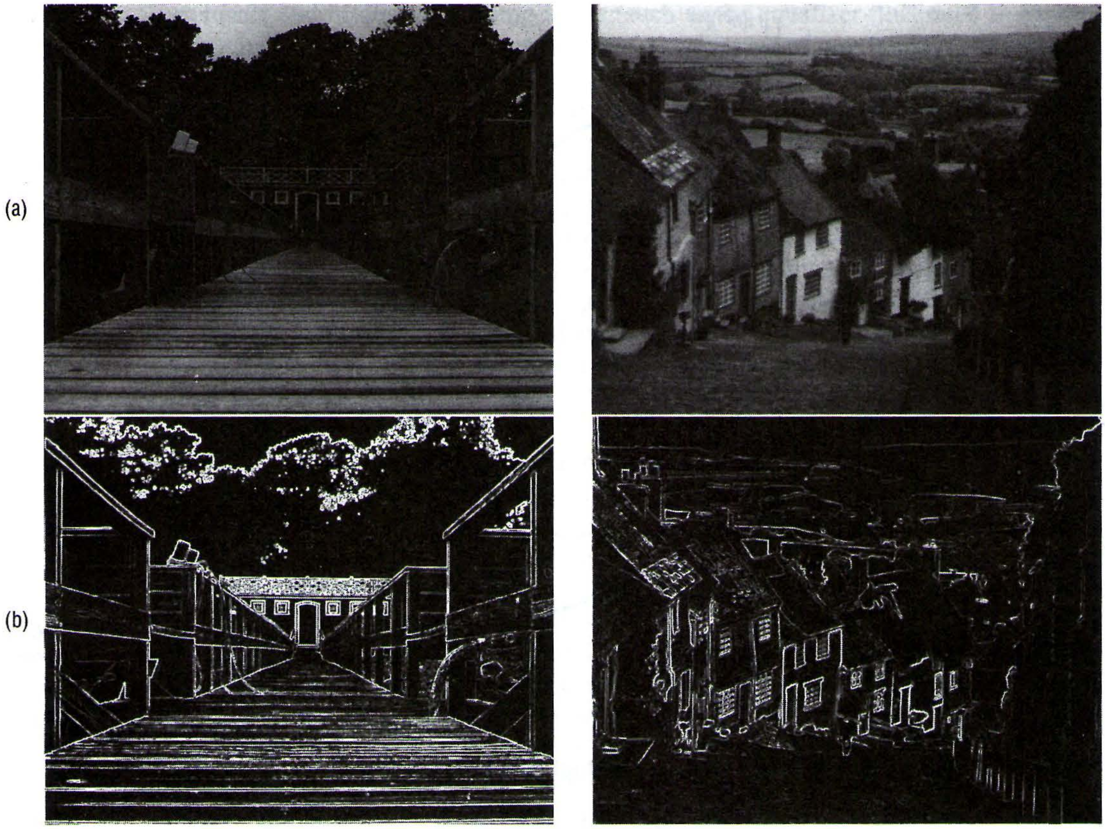


Figure 3.24 Effect of a differentiating filter on an image: (a) two original images; (b) the result of processing the original images with a differentiating filter.

In this case $h[n] = \frac{1}{2}(\delta[n] + \delta[n-1])$, and from eq. (3.122), we see that the frequency response of the system is

$$H(e^{j\omega}) = \frac{1}{2}[1 + e^{-j\omega}] = e^{-j\omega/2}\cos(\omega/2). \tag{3.139}$$

The magnitude of $H(e^{j\omega})$ is plotted in Figure 3.25(a), and $\sphericalangle H(e^{j\omega})$ is shown in Figure 3.25(b). As discussed in Section 1.3.3, low frequencies for discrete-time complex exponentials occur near $\omega = 0, \pm 2\pi, \pm 4\pi, \ldots$, and high frequencies near $\omega = \pm\pi, \pm 3\pi, \ldots$. This is a result of the fact that $e^{j(\omega+2\pi)n} = e^{j\omega n}$, so that in discrete time we need only consider a 2π interval of values of ω in order to cover a complete range of distinct discrete-time frequencies. As a consequence, any discrete-time frequency responses $H(e^{j\omega})$ must be periodic with period 2π, a fact that can also be deduced directly from eq. (3.122).

For the specific filter defined in eqs. (3.138) and (3.139), we see from Figure 3.25(a) that $|H(e^{j\omega})|$ is large for frequencies near $\omega = 0$ and decreases as we increase $|\omega|$ toward π, indicating that higher frequencies are attentuated more than lower ones. For example, if the input to this system is constant—i.e., a zero-frequency complex exponential

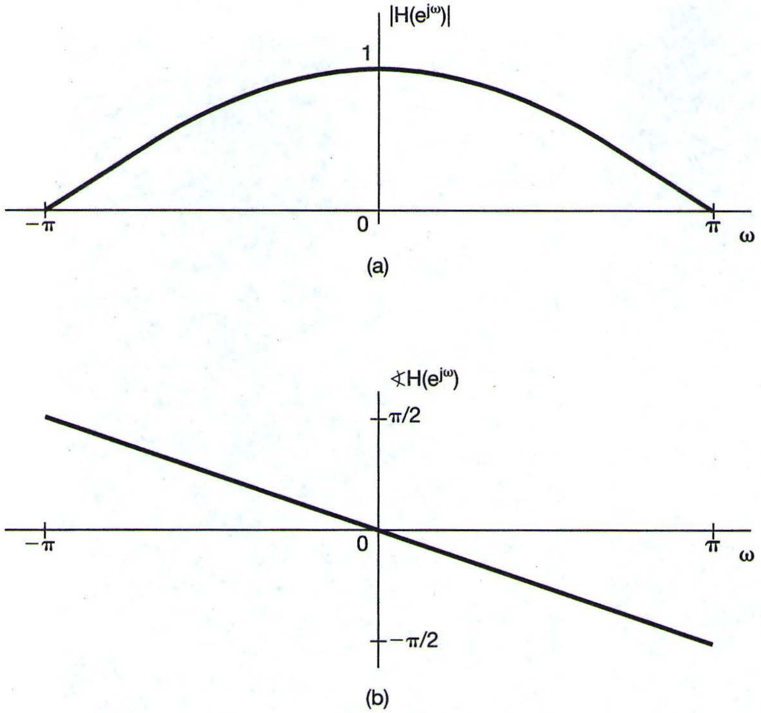


Figure 3.25 (a) Magnitude and (b) phase for the frequency response of the discrete-time LTI system $y[n] = 1/2(x[n] + x[n-1])$.

$x[n] = Ke^{j0\cdot n} = K$—then the output will be

$$y[n] = H(e^{j\cdot 0})Ke^{j\omega 0\cdot n} = K = x[n].$$

On the other hand, if the input is the high-frequency signal $x[n] = Ke^{j\pi n} = K(-1)^n$, then the output will be

$$y[n] = H(e^{j\pi})Ke^{j\pi\cdot n} = 0.$$

Thus, this system separates out the long-term constant value of a signal from its high-frequency fluctuations and, consequently, represents a first example of frequency-selective filtering, a topic we look at more carefully in the next subsection.

3.9.2 Frequency-Selective Filters

Frequency-selective filters are a class of filters specifically intended to accurately or approximately select some bands of frequencies and reject others. The use of frequency-selective filters arises in a variety of situations. For example, if noise in an audio recording is in a higher frequency band than the music or voice on the recording is, it can be removed by frequency-selective filtering. Another important application of frequency-selective filters is in communication systems. As we discuss in detail in Chapter 8, the basis for amplitude modulation (AM) systems is the transmission of information from many different sources simultaneously by putting the information from each channel into a separate frequency band and extracting the individual channels or bands at the receiver using frequency-selective filters. Frequency-selective filters for separating the individual channels and frequency-shaping filters (such as the equalizer illustrated in Figure 3.22) for adjusting the quality of the tone form a major part of any home radio and television receiver.

While frequency selectivity is not the only issue of concern in applications, its broad importance has led to a widely accepted set of terms describing the characteristics of frequency-selective filters. In particular, while the nature of the frequencies to be passed by a frequency-selective filter varies considerably from application to application, several basic types of filter are widely used and have been given names indicative of their function. For example, a *lowpass filter* is a filter that passes low frequencies—i.e., frequencies around $\omega = 0$—and attenuates or rejects higher frequencies. A *highpass filter* is a filter that passes high frequencies and attentuates or rejects low ones, and a *bandpass filter* is a filter that passes a band of frequencies and attenuates frequencies both higher and lower than those in the band that is passed. In each case, the *cutoff frequencies* are the frequencies defining the boundaries between frequencies that are passed and frequencies that are rejected—i.e., the frequencies in the *passband* and *stopband*.

Numerous questions arise in defining and assessing the quality of a frequency-selective filter. How effective is the filter at passing frequencies in the passband? How effective is it at attentuating frequencies in the stopband? How sharp is the transition near the cutoff frequency—i.e., from nearly free of distortion in the passband to highly attenuated in the stopband? Each of these questions involves a comparison of the characteristics of an actual frequency-selective filter with those of a filter with idealized behavior. Specifically, an *ideal frequency-selective filter* is a filter that exactly passes complex exponentials at one set of frequencies without any distortion and completely rejects signals at all other frequencies. For example, a continuous-time *ideal lowpass filter* with cutoff frequency ω_c is an LTI system that passes complex exponentials $e^{j\omega t}$ for values of ω in the range $-\omega_c \leq \omega \leq \omega_c$ and rejects signals at all other frequencies. That is, the frequency response of a continuous-time ideal lowpass filter is

$$H(j\omega) = \begin{cases} 1, & |\omega| \leq \omega_c \\ 0, & |\omega| > \omega_c \end{cases}, \tag{3.140}$$

as shown in Figure 3.26.

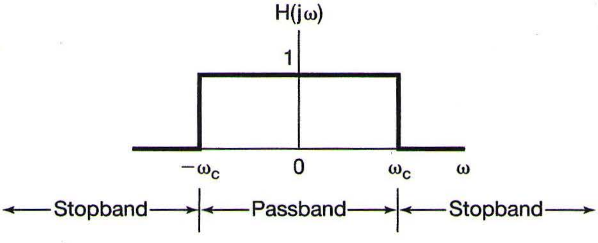


Figure 3.26 Frequency response of an ideal lowpass filter.

Figure 3.27(a) depicts the frequency response of an ideal continuous-time highpass filter with cutoff frequency ω_c, and Figure 3.27(b) illustrates an ideal continuous-time bandpass filter with lower cutoff frequency ω_{c1} and upper cutoff frequency ω_{c2}. Note that each of these filters is symmetric about $\omega = 0$, and thus, there appear to be two passbands for the highpass and bandpass filters. This is a consequence of our having adopted the

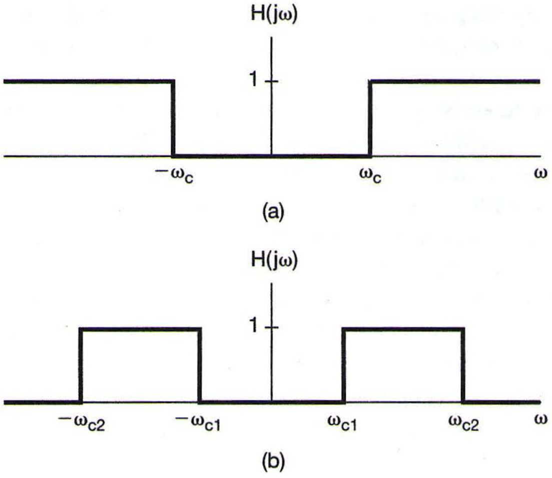

H(jω)
1
$-\omega_c$
ω_c
ω
(a)
H(jω)
1
$-\omega_{c2}$
$-\omega_{c1}$
ω_{c1}
ω_{c2}
ω
(b)

Figure 3.27 (a) Frequency response of an ideal highpass filter; (b) frequency response of an ideal bandpass filter.

use of the complex exponential signal $e^{j\omega t}$, rather than the sinusoidal signals $\sin \omega t$ and $\cos \omega t$, at frequency ω. Since $e^{j\omega t} = \cos \omega t + j \sin \omega t$ and $e^{-j\omega t} = \cos \omega t - j \sin \omega t$, both of these complex exponentials are composed of sinusoidal signals at the same frequency ω. For this reason, we usually define ideal filters so that they have the symmetric frequency response behavior seen in Figures 3.26 and 3.27.

In a similar fashion, we can define the corresponding set of ideal discrete-time frequency-selective filters, the frequency responses for which are depicted in Figure 3.28.

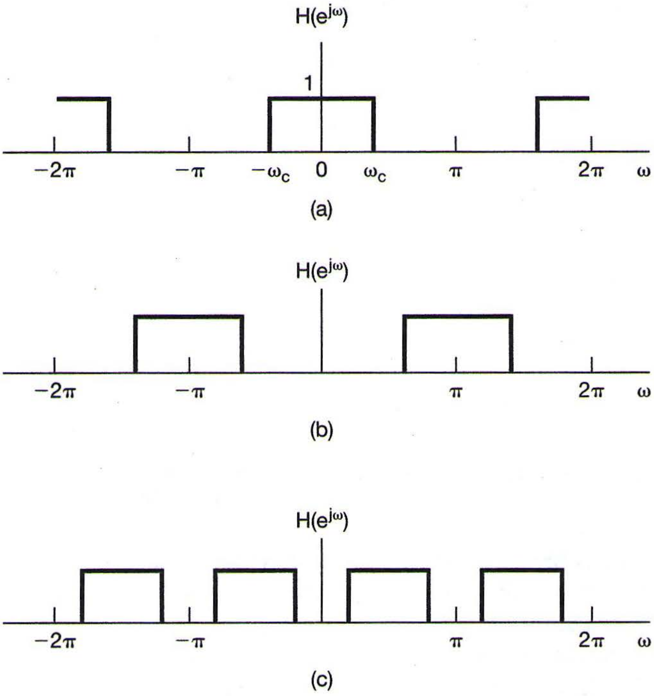

$H(e^{j\omega})$
1
-2π
$-\pi$
$-\omega_c$
0
ω_c
π
2π
ω
(a)
$H(e^{j\omega})$
-2π
$-\pi$
π
2π
ω
(b)
$H(e^{j\omega})$
-2π
$-\pi$
π
2π
ω
(c)

Figure 3.28 Discrete-time ideal frequency-selective filters: (a) lowpass; (b) highpass; (c) bandpass.

In particular, Figure 3.28(a) depicts an ideal discrete-time lowpass filter, Figure 3.28(b) is an ideal highpass filter, and Figure 3.28(c) is an ideal bandpass filter. Note that, as discussed in the preceding section, the characteristics of the continuous-time and discrete-time ideal filters differ by virtue of the fact that, for discrete-time filters, the frequency response $H(e^{j\omega})$ must be periodic with period 2π, with low frequencies near even multiples of π and high frequencies near odd multiples of π.

As we will see on numerous occasions, ideal filters are quite useful in describing idealized system configurations for a variety of applications. However, they are not realizable in practice and must be approximated. Furthermore, even if they could be realized, some of the characteristics of ideal filters might make them undesirable for particular applications, and a nonideal filter might in fact be preferable.

In detail, the topic of filtering encompasses many issues, including design and implementation. While we will not delve deeply into the details of filter design methodologies, in the remainder of this chapter and the following chapters we will see a number of other examples of both continuous-time and discrete-time filters and will develop the concepts and techniques that form the basis of this very important engineering discipline.

3.10 EXAMPLES OF CONTINUOUS-TIME FILTERS DESCRIBED BY DIFFERENTIAL EQUATIONS

In many applications, frequency-selective filtering is accomplished through the use of LTI systems described by linear constant-coefficient differential or difference equations. The reasons for this are numerous. For example, many physical systems that can be interpreted as performing filtering operations are characterized by differential or difference equations. A good example of this that we will examine in Chapter 6 is an automobile suspension system, which in part is designed to filter out high-frequency bumps and irregularities in road surfaces. A second reason for the use of filters described by differential or difference equations is that they are conveniently implemented using either analog or digital hardware. Furthermore, systems described by differential or difference equations offer an extremely broad and flexible range of designs, allowing one, for example, to produce filters that are close to ideal or that possess other desirable characteristics. In this and the next section, we consider several examples that illustrate the implementation of continuous-time and discrete-time frequency-selective filters through the use of differential and difference equations. In Chapters 4–6, we will see other examples of these classes of filters and will gain additional insights into the properties that make them so useful.

3.10.1 A Simple RC Lowpass Filter

Electrical circuits are widely used to implement continuous-time filtering operations. One of the simplest examples of such a circuit is the first-order *RC* circuit depicted in Figure 3.29, where the source voltage $v_s(t)$ is the system input. This circuit can be used to perform either a lowpass or highpass filtering operation, depending upon what we take as the output signal. In particular, suppose that we take the capacitor voltage $v_c(t)$ as the output. In this case, the output voltage is related to the input voltage through the linear

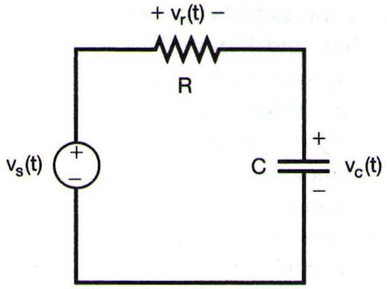


Figure 3.29 First-order *RC* filter.

constant-coefficient differential equation

$$RC\frac{dv_c(t)}{dt} + v_c(t) = v_s(t). \tag{3.141}$$

Assuming initial rest, the system described by eq. (3.141) is LTI. In order to determine its frequency response $H(j\omega)$, we note that, by definition, with input voltage $v_s(t) = e^{j\omega t}$, we must have the output voltage $v_c(t) = H(j\omega)e^{j\omega t}$. If we substitute these expressions into eq. (3.141), we obtain

$$RC\frac{d}{dt}[H(j\omega)e^{j\omega t}] + H(j\omega)e^{j\omega t} = e^{j\omega t}, \tag{3.142}$$

or

$$RC\,j\omega H(j\omega)e^{j\omega t} + H(j\omega)e^{j\omega t} = e^{j\omega t}, \tag{3.143}$$

from which it follows directly that

$$H(j\omega)e^{j\omega t} = \frac{1}{1 + RC\,j\omega}e^{j\omega t}, \tag{3.144}$$

or

$$H(j\omega) = \frac{1}{1 + RC\,j\omega}. \tag{3.145}$$

The magnitude and phase of the frequency response $H(j\omega)$ for this example are shown in Figure 3.30. Note that for frequencies near $\omega = 0$, $|H(j\omega)| \approx 1$, while for larger values of ω (positive or negative), $|H(j\omega)|$ is considerably smaller and in fact steadily decreases as $|\omega|$ increases. Thus, this simple *RC* filter (with $v_c(t)$ as output) is a nonideal lowpass filter.

To provide a first glimpse at the trade-offs involved in filter design, let us briefly consider the time-domain behavior of the circuit. In particular, the impulse response of the system described by eq. (3.141) is

$$h(t) = \frac{1}{RC}e^{-t/RC}u(t), \tag{3.146}$$

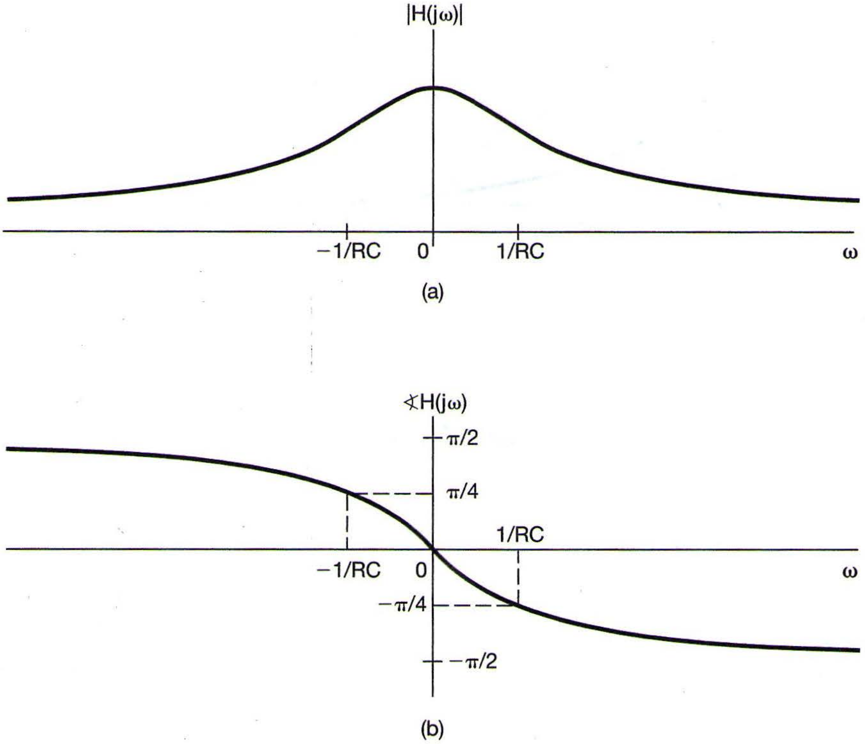


Figure 3.30 (a) Magnitude and (b) phase plots for the frequency response for the *RC* circuit of Figure 3.29 with output $v_c(t)$.

and the step response is

$$s(t) = [1 - e^{-t/RC}]u(t), \tag{3.147}$$

both of which are plotted in Figure 3.31 (where $\tau = RC$). Comparing Figures 3.30 and 3.31, we see a fundamental trade-off. Specifically, suppose that we would like our filter to pass only very low frequencies. From Figure 3.30(a), this implies that $1/RC$ must be small, or equivalently, that RC is large, so that frequencies other than the low ones of interest will be attentuated sufficiently. However, looking at Figure 3.31(b), we see that if RC is large, then the step response will take a considerable amount of time to reach its long-term value of 1. That is, the system responds sluggishly to the step input. Conversely, if we wish to have a faster step response, we need a smaller value of RC, which in turn implies that the filter will pass higher frequencies. This type of trade-off between behavior in the frequency domain and in the time domain is typical of the issues arising in the design and analysis of LTI systems and filters and is a subject we will look at more carefully in Chapter 6.

3.10.2 A Simple RC Highpass Filter

As an alternative to choosing the capacitor voltage as the output in our *RC* circuit, we can choose the voltage across the resistor. In this case, the differential equation relating input

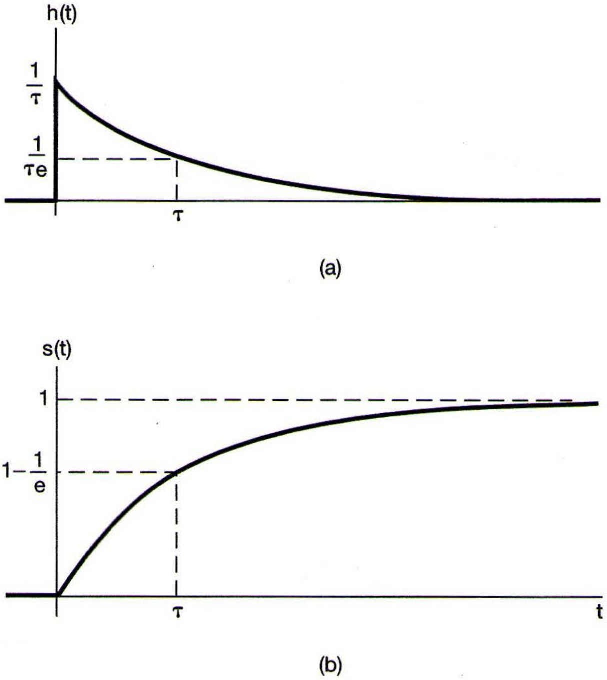


Figure 3.31 (a) Impulse response of the first-order *RC* lowpass filter with $\tau = RC$; (b) step response of *RC* lowpass filter with $\tau = RC$.

and output is

$$RC\frac{dv_r(t)}{dt} + v_r(t) = RC\frac{dv_s(t)}{dt}. \tag{3.148}$$

We can find the frequency response $G(j\omega)$ of this system in exactly the same way we did in the previous case: If $v_s(t) = e^{j\omega t}$, then we must have $v_r(t) = G(j\omega)e^{j\omega t}$; substituting these expressions into eq. (3.148) and performing a bit of algebra, we find that

$$G(j\omega) = \frac{j\omega RC}{1 + j\omega RC}. \tag{3.149}$$

The magnitude and phase of this frequency response are shown in Figure 3.32. From the figure, we see that the system attenuates lower frequencies and passes higher frequencies—i.e., those for which $|\omega| \gg 1/RC$—with minimal attenuation. That is, this system acts as a nonideal highpass filter.

As with the lowpass filter, the parameters of the circuit control both the frequency response of the highpass filter and its time response characteristics. For example, consider the step response for the filter. From Figure 3.29, we see that $v_r(t) = v_s(t) - v_c(t)$. Thus, if $v_s(t) = u(t)$, $v_c(t)$ must be given by eq. (3.147). Consequently, the step response of the highpass filter is

$$v_r(t) = e^{-t/RC}u(t), \tag{3.150}$$

which is depicted in Figure 3.33. Consequently, as *RC* is increased, the response becomes more sluggish—i.e., the step response takes a longer time to reach its long-term value

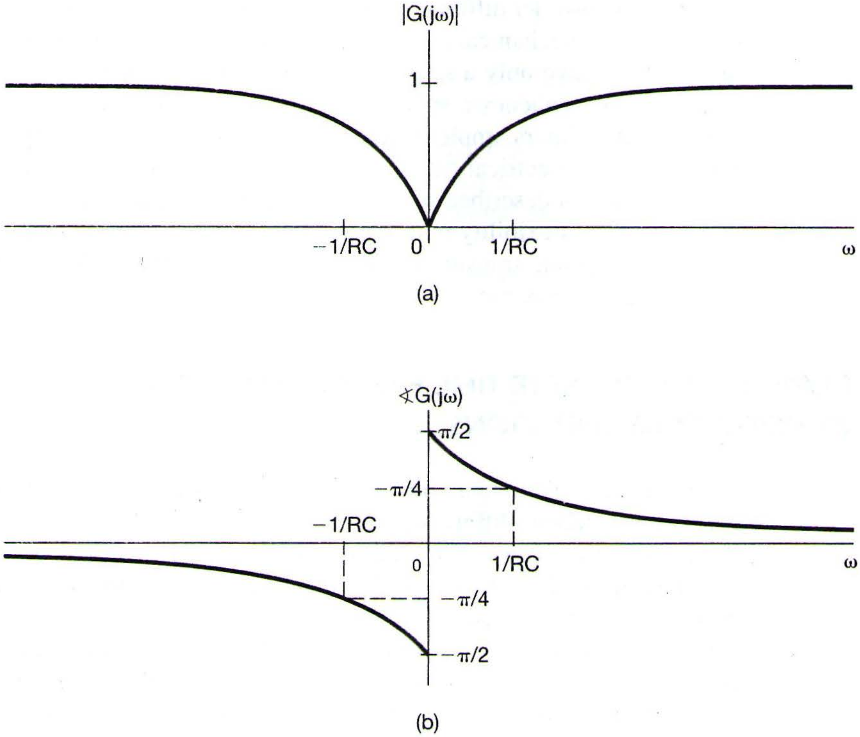


Figure 3.32 (a) Magnitude and (b) phase plots for the frequency response of the *RC* circuit of Figure 3.29 with output $v_r(t)$.

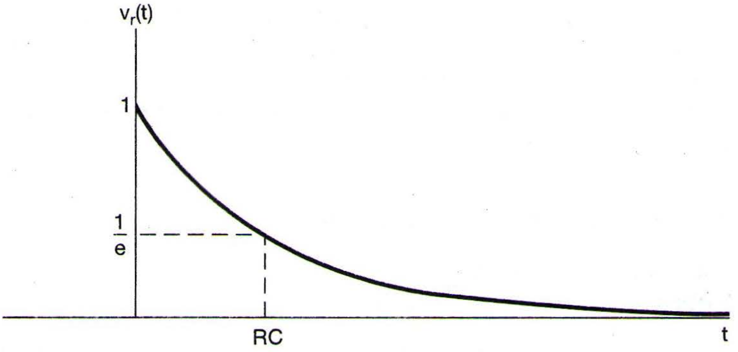


Figure 3.33 Step response of the first-order *RC* highpass filter with $\tau = RC$.

of 0. From Figure 3.32, we see that increasing *RC* (or equivalently, decreasing $1/RC$) also affects the frequency response, specifically, it extends the passband down to lower frequencies.

We observe from the two examples in this section that a simple *RC* circuit can serve as a rough approximation to a highpass or a lowpass filter, depending upon the choice of the physical output variable. As illustrated in Problem 3.71, a simple mechanical system using a mass and a mechanical damper can also serve as a lowpass or highpass filter described by analogous first-order differential equations. Because of their simplicity, these examples of electrical and mechanical filters do not have a sharp transition from passband to stopband and, in fact, have only a single parameter (namely, *RC* in the electrical case) that controls both the frequency response and time response behavior of the system. By designing more complex filters, implemented using more energy storage elements (capacitances and inductances in electrical filters and springs and damping devices in mechanical filters), we obtain filters described by higher order differential equations. Such filters offer considerably more flexibility in terms of their characteristics, allowing, for example, sharper passband-stopband transition or more control over the trade-offs between time response and frequency response.

3.11 EXAMPLES OF DISCRETE-TIME FILTERS DESCRIBED BY DIFFERENCE EQUATIONS

As with their continuous-time counterparts, discrete-time filters described by linear constant-coefficient difference equations are of considerable importance in practice. Indeed, since they can be efficiently implemented in special- or general-purpose digital systems, filters described by difference equations are widely used in practice. As in almost all aspects of signal and system analysis, when we examine discrete-time filters described by difference equations, we find both strong similarities and important differences with the continuous-time case. In particular, discrete-time LTI systems described by difference equations can either be recursive and have impulse responses of infinite length (IIR systems) or be nonrecursive and have finite-length impulse responses (FIR systems). The former are the direct counterparts of continuous-time systems described by differential equations illustrated in the previous section, while the latter are also of considerable practical importance in digital systems. These two classes have distinct sets of advantages and disadvantages in terms of ease of implementation and in terms of the order of filter or the complexity required to achieve particular design objectives. In this section we limit ourselves to a few simple examples of recursive and nonrecursive filters, while in Chapters 5 and 6 we develop additional tools and insights that allow us to analyze and understand the properties of these systems in more detail.

3.11.1 First-Order Recursive Discrete-Time Filters

The discrete-time counterpart of each of the first-order filters considered in Section 3.10 is the LTI system described by the first-order difference equation

$$y[n] - ay[n-1] = x[n]. \tag{3.151}$$

From the eigenfunction property of complex exponential signals, we know that if $x[n] = e^{j\omega n}$, then $y[n] = H(e^{j\omega})e^{j\omega n}$, where $H(e^{j\omega})$ is the frequency response of the system. Substituting into eq. (3.151), we obtain

$$H(e^{j\omega})e^{j\omega n} - aH(e^{j\omega})e^{j\omega(n-1)} = e^{j\omega n}, \tag{3.152}$$

or

$$[1 - ae^{-j\omega}]H(e^{j\omega})e^{j\omega n} = e^{j\omega n}, \tag{3.153}$$

so that

$$H(e^{j\omega}) = \frac{1}{1 - ae^{-j\omega}}. \tag{3.154}$$

The magnitude and phase of $H(e^{j\omega})$ are shown in Figure 3.34(a) for $a = 0.6$ and in Figure 3.34(b) for $a = -0.6$. We observe that, for the positive value of a, the difference equation (3.151) behaves like a lowpass filter with minimal attenuation of low frequencies near $\omega = 0$ and increasing attenuation as we increase ω toward $\omega = \pi$. For the negative value of a, the system is a highpass filter, passing frequencies near $\omega = \pi$ and attenuating lower frequencies. In fact, for any positive value of $a < 1$, the system approximates a lowpass filter, and for any negative value of $a > -1$, the system approximates a highpass filter, where $|a|$ controls the size of the filter passband, with broader passbands as $|a|$ is decreased.

As with the continuous-time examples, we again have a trade-off between time domain and frequency domain characteristics. In particular, the impulse response of the system described by eq. (3.151) is

$$h[n] = a^n u[n]. \tag{3.155}$$

The step response $s[n] = u[n] * h[n]$ is

$$s[n] = \frac{1 - a^{n+1}}{1 - a} u[n]. \tag{3.156}$$

From these expressions, we see that $|a|$ also controls the speed with which the impulse and step responses approach their long-term values, with faster responses for smaller values of $|a|$, and hence, for filters with smaller passbands. Just as with differential equations, higher order recursive difference equations can be used to provide sharper filter characteristics and to provide more flexibility in balancing time-domain and frequency-domain constraints.

Finally, note from eq. (3.155) that the system described by eq. (3.151) is unstable if $|a| \geq 1$ and thus does not have a finite response to complex exponential inputs. As we stated previously, Fourier-based methods and frequency domain analysis focus on systems with finite responses to complex exponentials; hence, for examples such as eq. (3.151), we restrict ourselves to stable systems.

3.11.2 Nonrecursive Discrete-Time Filters

The general form of an FIR nonrecursive difference equation is

$$y[n] = \sum_{k=-N}^{M} b_k x[n - k]. \tag{3.157}$$

That is, the output $y[n]$ is a *weighted average* of the $(N + M + 1)$ values of $x[n]$ from $x[n - M]$ through $x[n + N]$, with the weights given by the coefficients b_k. Systems of this form can be used to meet a broad array of filtering objectives, including frequency-selective filtering.

One frequently used example of such a filter is a *moving-average filter,* where the output $y[n]$ for any n—say, n_0—is an average of values of $x[n]$ in the vicinity of n_0. The

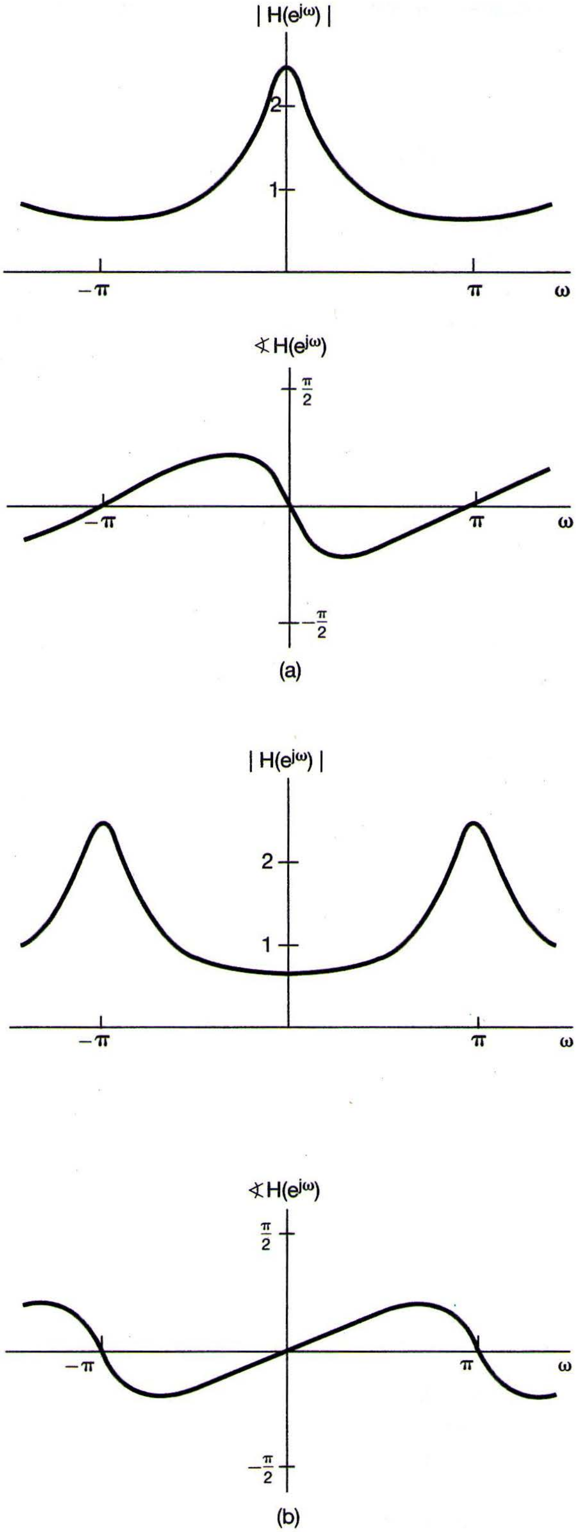


Figure 3.34 Frequency response of the first-order recursive discrete-time filter of eq. (3.151): (a) $a = 0.6$; (b) $a = -0.6$.

basic idea is that by averaging values locally, rapid high-frequency components of the input will be averaged out and lower frequency variations will be retained, corresponding to smoothing or lowpass filtering the original sequence. A simple two-point moving-average filter was briefly introduced in Section 3.9 [eq. (3.138)]. An only slightly more complex example is the three-point moving-average filter, which is of the form

$$y[n] = \frac{1}{3}(x[n-1] + x[n] + x[n+1]), \tag{3.158}$$

so that each output $y[n]$ is the average of three consecutive input values. In this case,

$$h[n] = \frac{1}{3}[\delta[n+1] + \delta[n] + \delta[n-1]],$$

and thus, from eq. (3.122), the corresponding frequency response is

$$H(e^{j\omega}) = \frac{1}{3}[e^{j\omega} + 1 + e^{-j\omega}] = \frac{1}{3}(1 + 2\cos\omega). \tag{3.159}$$

The magnitude of $H(e^{j\omega})$ is sketched in Figure 3.35. We observe that the filter has the general characteristics of a lowpass filter, although, as with the first-order recursive filter, it does not have a sharp transition from passband to stopband.

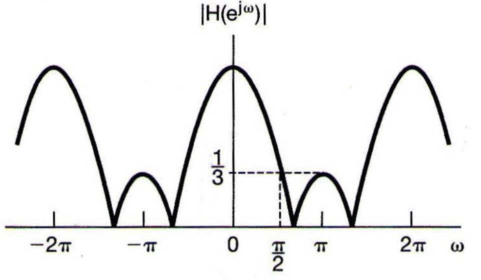


Figure 3.35 Magnitude of the frequency response of a three-point moving-average lowpass filter.

The three-point moving-average filter in eq. (3.158) has no parameters that can be changed to adjust the effective cutoff frequency. As a generalization of this moving-average filter, we can consider averaging over $N + M + 1$ neighboring points—that is, using a difference equation of the form

$$y[n] = \frac{1}{N+M+1}\sum_{k=-N}^{M} x[n-k]. \tag{3.160}$$

The corresponding impulse response is a rectangular pulse (i.e., $h[n] = 1/(N + M + 1)$ for $-N \leq n \leq M$, and $h[n] = 0$ otherwise). The filter's frequency response is

$$H(e^{j\omega}) = \frac{1}{N+M+1}\sum_{k=-N}^{M} e^{-j\omega k}. \tag{3.161}$$

The summation in eq. (3.161) can be evaluated by performing calculations similar to those in Example 3.12, yielding

$$H(e^{j\omega}) = \frac{1}{N+M+1} e^{j\omega[(N-M)/2]} \frac{\sin[\omega(M+N+1)/2]}{\sin(\omega/2)}. \tag{3.162}$$

By adjusting the size, $N + M + 1$, of the averaging window we can vary the cutoff frequency. For example, the magnitude of $H(e^{j\omega})$ is shown in Figure 3.36 for $M+N+1 = 33$ and $M + N + 1 = 65$.

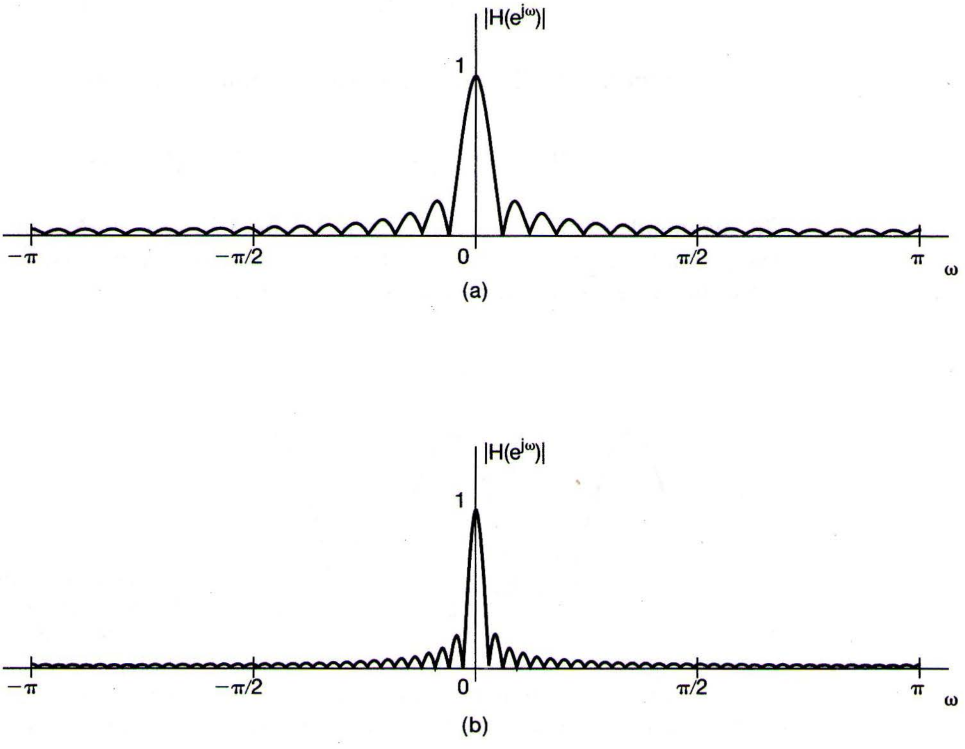


Figure 3.36 Magnitude of the frequency response for the lowpass moving-average filter of eq. (3.162): (a) $M = N = 16$; (b) $M = N = 32$.

Nonrecursive filters can also be used to perform highpass filtering operations. To illustrate this, again with a simple example, consider the difference equation

$$y[n] = \frac{x[n] - x[n-1]}{2}. \tag{3.163}$$

For input signals that are approximately constant, the value of $y[n]$ is close to zero. For input signals that vary greatly from sample to sample, the values of $y[n]$ can be expected to have larger amplitude. Thus, the system described by eq. (3.163) approximates a highpass filtering operation, attenuating slowly varying low-frequency components and passing rapidly varying higher frequency components with little attenuation. To see this more precisely we need to look at the system's frequency response. In this case, $h[n] = \frac{1}{2}\{\delta[n] - \delta[n-1]\}$, so that direct application of eq. (3.122) yields

$$H(e^{j\omega}) = \frac{1}{2}[1 - e^{-j\omega}] = je^{j\omega/2}\sin(\omega/2). \tag{3.164}$$

In Figure 3.37 we have plotted the magnitude of $H(e^{j\omega})$, showing that this simple system approximates a highpass filter, albeit one with a very gradual transition from passband to stopband. By considering more general nonrecursive filters, we can achieve far sharper transitions in lowpass, highpass, and other frequency-selective filters.

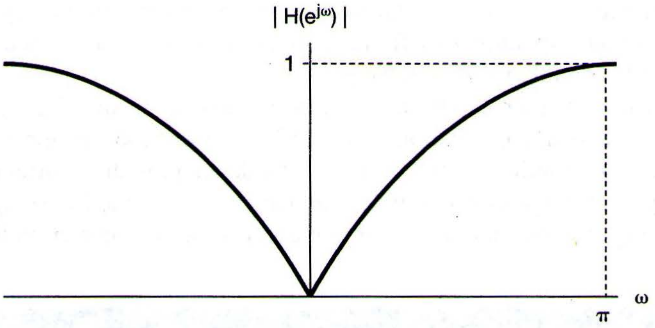


Figure 3.37 Frequency response of a simple highpass filter.

Note that, since the impulse response of any FIR system is of finite length (i.e., from eq. (3.157), $h[n] = b_n$ for $-N \le n \le M$ and 0 otherwise), it is always absolutely summable for any choices of the b_n. Hence, all such filters are stable. Also, if $N > 0$ in eq. (3.157), the system is noncausal, since $y[n]$ then depends on future values of the input. In some applications, such as those involving the processing of previously recorded signals, causality is not a necessary constraint, and thus, we are free to use filters with $N > 0$. In others, such as many involving real-time processing, causality is essential, and in such cases we must take $N \le 0$.

3.12 SUMMARY

In this chapter, we have introduced and developed Fourier series representations for both continuous-time and discrete-time systems and have used these representations to take a first look at one of the very important applications of the methods of signal and system analysis, namely, filtering. In particular, as we discussed in Section 3.2, one of the primary motivations for the use of Fourier series is the fact that complex exponential signals are eigenfunctions of LTI systems. We have also seen, in Sections 3.3–3.7, that any periodic signal of practical interest can be represented in a Fourier series—i.e., as a weighted sum of harmonically related complex exponentials that share a common period with the signal being represented. In addition, we have seen that the Fourier series representation has a number of important properties which describe how different characteristics of signals are reflected in their Fourier series coefficients.

One of the most important properties of Fourier series is a direct consequence of the eigenfunction property of complex exponentials. Specifically, if a periodic signal is applied to an LTI system, then the output will be periodic with the same period, and each of the Fourier coefficients of the output is the corresponding Fourier coefficient of the input multiplied by a complex number whose value is a function of the frequency corresponding to that Fourier coefficient. This function of frequency is characteristic of the LTI system and is referred to as the frequency response of the system. By examining the frequency response, we were led directly to the idea of filtering of signals using LTI systems, a concept that has numerous applications, including several that we have described. One important class of applications involves the notion of frequency-selective filtering—i.e., the idea of using an LTI system to pass certain specified bands of frequencies and stop or significantly attentuate others. We introduced the concept of ideal frequency-selective filters and also gave several examples of frequency-selective filters described by linear constant-coefficient differential or difference equations.

The purpose of this chapter has been to begin the process of developing both the tools of Fourier analysis and an appreciation for the utility of these tools in applications. In the chapters that follow, we continue with this agenda by developing the Fourier transform representations for aperiodic signals in continuous and discrete time and by taking a deeper look not only at filtering, but also at other important applications of Fourier methods.

Chapter 3 Problems

The first section of problems belongs to the basic category and the answers are provided in the back of the book. The remaining three sections contain problems belonging to the basic, advanced, and extension categories, respectively.

BASIC PROBLEMS WITH ANSWERS

3.1. A continuous-time periodic signal $x(t)$ is real valued and has a fundamental period $T = 8$. The nonzero Fourier series coefficients for $x(t)$ are

$$a_1 = a_{-1} = 2, a_3 = a_{-3}^* = 4j.$$

Express $x(t)$ in the form

$$x(t) = \sum_{k=0}^{\infty} A_k \cos(\omega_k t + \phi_k).$$

3.2. A discrete-time periodic signal $x[n]$ is real valued and has a fundamental period $N = 5$. The nonzero Fourier series coefficients for $x[n]$ are

$$a_0 = 1, a_2 = a_{-2}^* = e^{j\pi/4}, a_4 = a_{-4}^* = 2e^{j\pi/3}.$$

Express $x[n]$ in the form

$$x[n] = A_0 + \sum_{k=1}^{\infty} A_k \sin(\omega_k n + \phi_k).$$

3.3. For the continuous-time periodic signal

$$x(t) = 2 + \cos\left(\frac{2\pi}{3}t\right) + 4\sin\left(\frac{5\pi}{3}t\right),$$

determine the fundamental frequency ω_0 and the Fourier series coefficients a_k such that

$$x(t) = \sum_{k=-\infty}^{\infty} a_k e^{jk\omega_0 t}.$$

3.4. Use the Fourier series analysis equation (3.39) to calculate the coefficients a_k for the continuous-time periodic signal

$$x(t) = \begin{cases} 1.5, & 0 \le t < 1 \\ -1.5, & 1 \le t < 2 \end{cases}$$

with fundamental frequency $\omega_0 = \pi$.

3.5. Let $x_1(t)$ be a continuous-time periodic signal with fundamental frequency ω_1 and Fourier coefficients a_k. Given that

$$x_2(t) = x_1(1-t) + x_1(t-1),$$

how is the fundamental frequency ω_2 of $x_2(t)$ related to ω_1? Also, find a relationship between the Fourier series coefficients b_k of $x_2(t)$ and the coefficients a_k. You may use the properties listed in Table 3.1.

3.6. Consider three continuous-time periodic signals whose Fourier series representations are as follows:

$$x_1(t) = \sum_{k=0}^{100} \left(\frac{1}{2}\right)^k e^{jk\frac{2\pi}{50}t},$$

$$x_2(t) = \sum_{k=-100}^{100} \cos(k\pi) e^{jk\frac{2\pi}{50}t},$$

$$x_3(t) = \sum_{k=-100}^{100} j\sin\left(\frac{k\pi}{2}\right) e^{jk\frac{2\pi}{50}t}.$$

Use Fourier series properties to help answer the following questions:

(a) Which of the three signals is/are real valued?

(b) Which of the three signals is/are even?

3.7. Suppose the periodic signal $x(t)$ has fundamental period T and Fourier coefficients a_k. In a variety of situations, it is easier to calculate the Fourier series coefficients

b_k for $g(t) = dx(t)/dt$, as opposed to calculating a_k directly. Given that

$$\int_T^{2T} x(t)\,dt = 2,$$

find an expression for a_k in terms of b_k and T. You may use any of the properties listed in Table 3.1 to help find the expression.

3.8. Suppose we are given the following information about a signal $x(t)$:

1. $x(t)$ is real and odd.
2. $x(t)$ is periodic with period $T = 2$ and has Fourier coefficients a_k.
3. $a_k = 0$ for $|k| > 1$.
4. $\frac{1}{2}\int_0^2 |x(t)|^2\,dt = 1$.

Specify two different signals that satisfy these conditions.

3.9. Use the analysis equation (3.95) to evaluate the numerical values of one period of the Fourier series coefficients of the periodic signal

$$x[n] = \sum_{m=-\infty}^{\infty} \{4\delta[n-4m] + 8\delta[n-1-4m]\}.$$

3.10. Let $x[n]$ be a real and odd periodic signal with period $N = 7$ and Fourier coefficients a_k. Given that

$$a_{15} = j,\ a_{16} = 2j,\ a_{17} = 3j,$$

determine the values of a_0, a_{-1}, a_{-2}, and a_{-3}.

3.11. Suppose we are given the following information about a signal $x[n]$:

1. $x[n]$ is a real and even signal.
2. $x[n]$ has period $N = 10$ and Fourier coefficients a_k.
3. $a_{11} = 5$.
4. $\frac{1}{10}\sum_{n=0}^{9} |x[n]|^2 = 50$.

Show that $x[n] = A\cos(Bn + C)$, and specify numerical values for the constants A, B, and C.

3.12. Each of the two sequences $x_1[n]$ and $x_2[n]$ has a period $N = 4$, and the corresponding Fourier series coefficients are specified as

$$x_1[n] \longleftrightarrow a_k, \quad x_2[n] \longleftrightarrow b_k,$$

where

$$a_0 = a_3 = \frac{1}{2}a_1 = \frac{1}{2}a_2 = 1 \quad \text{and} \quad b_0 = b_1 = b_2 = b_3 = 1.$$

Using the multiplication property in Table 3.1, determine the Fourier series coefficients c_k for the signal $g[n] = x_1[n]x_2[n]$.

3.13. Consider a continuous-time LTI system whose frequency response is

$$H(j\omega) = \int_{-\infty}^{\infty} h(t)e^{-j\omega t}\,dt = \frac{\sin(4\omega)}{\omega}.$$

If the input to this system is a periodic signal

$$x(t) = \begin{cases} 1, & 0 \le t < 4 \\ -1, & 4 \le t < 8 \end{cases}$$

with period $T = 8$, determine the corresponding system output $y(t)$.

3.14. When the impulse train

$$x[n] = \sum_{k=-\infty}^{\infty} \delta[n-4k]$$

is the input to a particular LTI system with frequency response $H(e^{j\omega})$, the output of the system is found to be

$$y[n] = \cos\left(\frac{5\pi}{2}n + \frac{\pi}{4}\right).$$

Determine the values of $H(e^{jk\pi/2})$ for $k = 0, 1, 2$, and 3.

3.15. Consider a continuous-time ideal lowpass filter S whose frequency response is

$$H(j\omega) = \begin{cases} 1, & |\omega| \le 100 \\ 0, & |\omega| > 100 \end{cases}.$$

When the input to this filter is a signal $x(t)$ with fundamental period $T = \pi/6$ and Fourier series coefficients a_k, it is found that

$$x(t) \xrightarrow{S} y(t) = x(t).$$

For what values of k is it guaranteed that $a_k = 0$?

3.16. Determine the output of the filter shown in Figure P3.16 for the following periodic inputs:

(a) $x_1[n] = (-1)^n$

(b) $x_2[n] = 1 + \sin\left(\frac{3\pi}{8}n + \frac{\pi}{4}\right)$

(c) $x_3[n] = \sum_{k=-\infty}^{\infty}\left(\frac{1}{2}\right)^{n-4k} u[n-4k]$

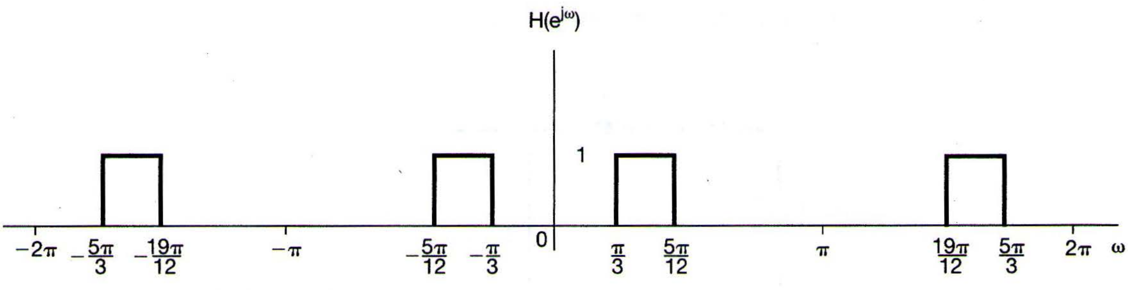


Figure P3.16

3.17. Consider three continuous-time systems S_1, S_2, and S_3 whose responses to a complex exponential input e^{j5t} are specified as

$$\begin{aligned} S_1 &: e^{j5t} \longrightarrow te^{j5t}, \\ S_2 &: e^{j5t} \longrightarrow e^{j5(t-1)}, \\ S_3 &: e^{j5t} \longrightarrow \cos(5t). \end{aligned}$$

For each system, determine whether the given information is sufficient to conclude that the system is definitely *not* LTI.

3.18. Consider three discrete-time systems S_1, S_2, and S_3 whose respective responses to a complex exponential input $e^{j\pi n/2}$ are specified as

$$\begin{aligned} S_1 &: e^{j\pi n/2} \longrightarrow e^{j\pi n/2}u[n], \\ S_2 &: e^{j\pi n/2} \longrightarrow e^{j3\pi n/2}, \\ S_3 &: e^{j\pi n/2} \longrightarrow 2e^{j5\pi n/2}. \end{aligned}$$

For each system, determine whether the given information is sufficient to conclude that the system is definitely *not* LTI.

3.19. Consider a causal LTI system implemented as the *RL* circuit shown in Figure P3.19. A current source produces an input current $x(t)$, and the system output is considered to be the current $y(t)$ flowing through the inductor.

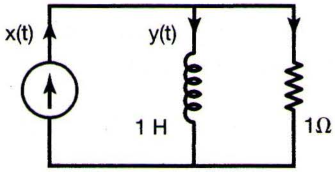


Figure P3.19

(a) Find the differential equation relating $x(t)$ and $y(t)$.
(b) Determine the frequency response of this system by considering the output of the system to inputs of the form $x(t) = e^{j\omega t}$.
(c) Determine the output $y(t)$ if $x(t) = \cos(t)$.

3.20. Consider a causal LTI system implemented as the *RLC* circuit shown in Figure P3.20. In this circuit, $x(t)$ is the input voltage. The voltage $y(t)$ across the capacitor is considered the system output.

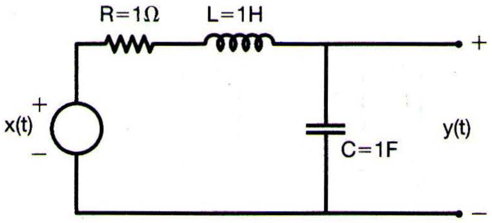


Figure P3.20

(a) Find the differential equation relating $x(t)$ and $y(t)$.
(b) Determine the frequency response of this system by considering the output of the system to inputs of the form $x(t) = e^{j\omega t}$.
(c) Determine the output $y(t)$ if $x(t) = \sin(t)$.

BASIC PROBLEMS

3.21. A continuous-time periodic signal $x(t)$ is real valued and has a fundamental period $T = 8$. The nonzero Fourier series coefficients for $x(t)$ are specified as

$$a_1 = a^*_{-1} = j, a_5 = a_{-5} = 2.$$

Express $x(t)$ in the form

$$x(t) = \sum_{k=0}^{\infty} A_k \cos(w_k t + \phi_k).$$

3.22. Determine the Fourier series representations for the following signals:
(a) Each $x(t)$ illustrated in Figure P3.22(a)–(f).
(b) $x(t)$ periodic with period 2 and

$$x(t) = e^{-t} \quad \text{for} \quad -1 < t < 1$$

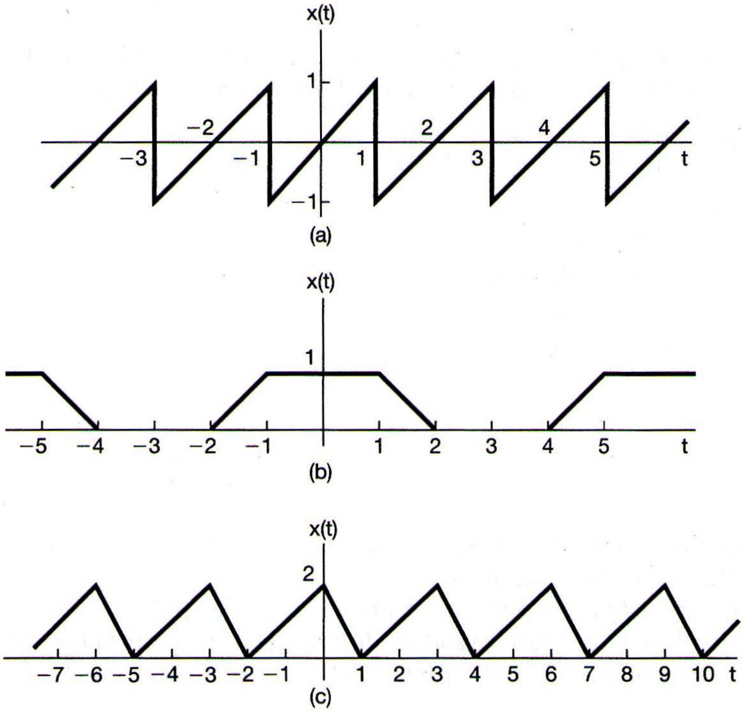


Figure P3.22

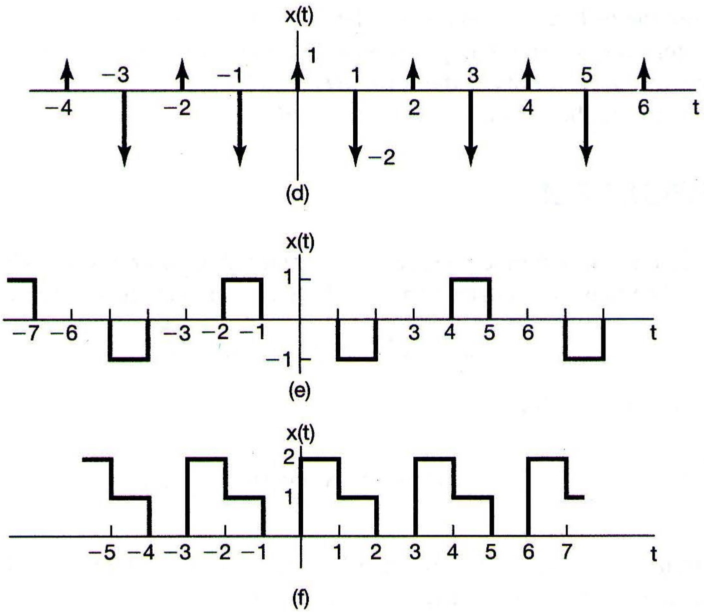


Figure P3.22 *Continued*

(c) $x(t)$ periodic with period 4 and

$$x(t) = \begin{cases} \sin \pi t, & 0 \leq t \leq 2 \\ 0, & 2 < t \leq 4 \end{cases}$$

3.23. In each of the following, we specify the Fourier series coefficients of a continuous-time signal that is periodic with period 4. Determine the signal $x(t)$ in each case.

(a) $a_k = \begin{cases} 0, & k = 0 \\ (j)^k \frac{\sin k\pi/4}{k\pi}, & \text{otherwise} \end{cases}$

(b) $a_k = (-1)^k \frac{\sin k\pi/8}{2k\pi}, \quad a_0 = \frac{1}{16}$

(c) $a_k = \begin{cases} jk, & |k| < 3 \\ 0, & \text{otherwise} \end{cases}$

(d) $a_k = \begin{cases} 1, & k \text{ even} \\ 2, & k \text{ odd} \end{cases}$

3.24. Let

$$x(t) = \begin{cases} t, & 0 \leq t \leq 1 \\ 2 - t, & 1 \leq t \leq 2 \end{cases}$$

be a periodic signal with fundamental period $T = 2$ and Fourier coefficients a_k.

(a) Determine the value of a_0.

(b) Determine the Fourier series representation of $dx(t)/dt$.

(c) Use the result of part (b) and the differentiation property of the continuous-time Fourier series to help determine the Fourier series coefficients of $x(t)$.

3.25. Consider the following three continuous-time signals with a fundamental period of $T = 1/2$:

$$\begin{aligned} x(t) &= \cos(4\pi t), \\ y(t) &= \sin(4\pi t), \\ z(t) &= x(t)y(t). \end{aligned}$$

(a) Determine the Fourier series coefficients of $x(t)$.
(b) Determine the Fourier series coefficients of $y(t)$.
(c) Use the results of parts (a) and (b), along with the multiplication property of the continuous-time Fourier series, to determine the Fourier series coefficients of $z(t) = x(t)y(t)$.
(d) Determine the Fourier series coefficients of $z(t)$ through direct expansion of $z(t)$ in trigonometric form, and compare your result with that of part (c).

3.26. Let $x(t)$ be a periodic signal whose Fourier series coefficients are

$$a_k = \begin{cases} 2, & k = 0 \\ j(\frac{1}{2})^{|k|}, & \text{otherwise} \end{cases}.$$

Use Fourier series properties to answer the following questions:
(a) Is $x(t)$ real?
(b) Is $x(t)$ even?
(c) Is $dx(t)/dt$ even?

3.27. A discrete-time periodic signal $x[n]$ is real valued and has a fundamental period $N = 5$. The nonzero Fourier series coefficients for $x[n]$ are

$$a_0 = 2, a_2 = a_{-2}^* = 2e^{j\pi/6}, \quad a_4 = a_{-4}^* = e^{j\frac{\pi}{3}}.$$

Express $x[n]$ in the form

$$x[n] = A_0 + \sum_{k=1}^{\infty} A_k \sin(\omega_k n + \phi_k).$$

3.28. Determine the Fourier series coefficients for each of the following discrete-time periodic signals. Plot the magnitude and phase of each set of coefficients a_k.
(a) Each $x[n]$ depicted in Figure P3.28(a)–(c)
(b) $x[n] = \sin(2\pi n/3)\cos(\pi n/2)$
(c) $x[n]$ periodic with period 4 and

$$x[n] = 1 - \sin\frac{\pi n}{4} \quad \text{for } 0 \leq n \leq 3$$

(d) $x[n]$ periodic with period 12 and

$$x[n] = 1 - \sin\frac{\pi n}{4} \quad \text{for } 0 \leq n \leq 11$$

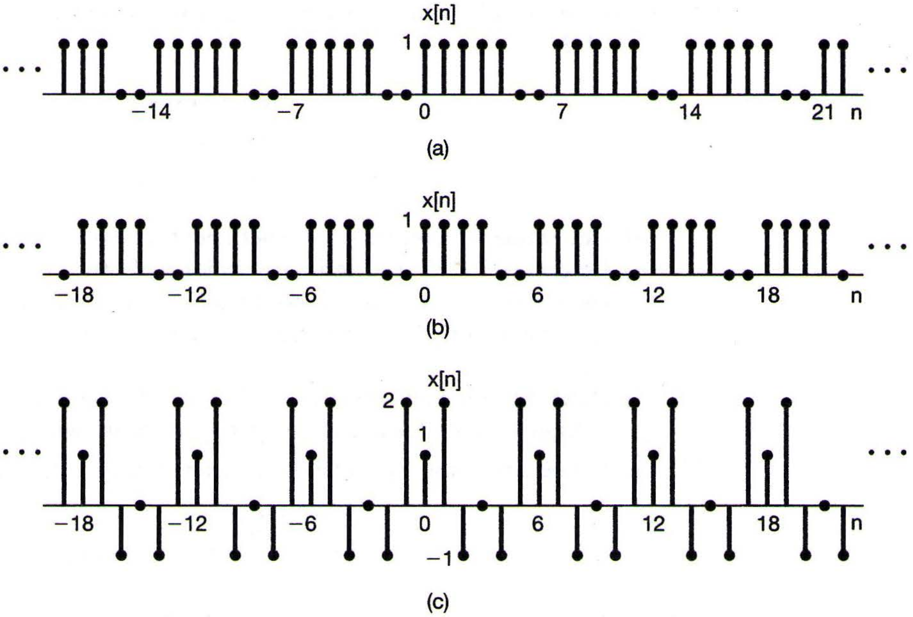


Figure P3.28

3.29. In each of the following, we specify the Fourier series coefficients of a signal that is periodic with period 8. Determine the signal $x[n]$ in each case.

(a) $a_k = \cos\left(\frac{k\pi}{4}\right) + \sin\left(\frac{3k\pi}{4}\right)$ **(b)** $a_k = \begin{cases} \sin\left(\frac{k\pi}{3}\right), & 0 \le k \le 6 \\ 0, & k = 7 \end{cases}$

(c) a_k as in Figure P3.29(a) **(d)** a_k as in Figure P3.29(b)

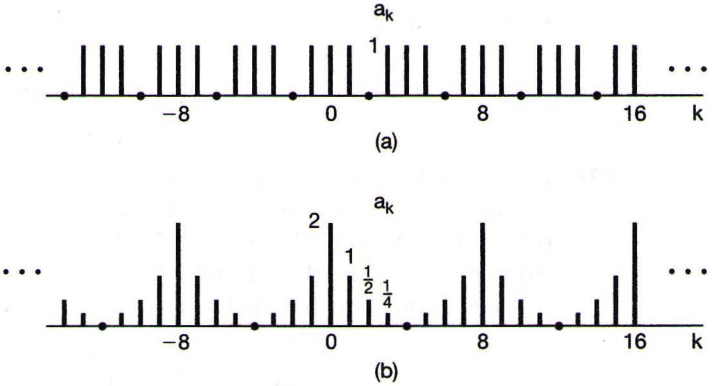


Figure P3.29

3.30. Consider the following three discrete-time signals with a fundamental period of 6:

$$x[n] = 1 + \cos\left(\frac{2\pi}{6}n\right), \qquad y[n] = \sin\left(\frac{2\pi}{6}n + \frac{\pi}{4}\right), \qquad z[n] = x[n]y[n].$$

(a) Determine the Fourier series coefficients of $x[n]$.
(b) Determine the Fourier series coefficients of $y[n]$.
(c) Use the results of parts (a) and (b), along with the multiplication property of the discrete-time Fourier series, to determine the Fourier series coefficients of $z[n] = x[n]y[n]$.
(d) Determine the Fourier series coefficients of $z[n]$ through direct evaluation, and compare your result with that of part (c).

3.31. Let

$$x[n] = \begin{cases} 1, & 0 \le n \le 7 \\ 0, & 8 \le n \le 9 \end{cases}$$

be a periodic signal with fundamental period $N = 10$ and Fourier series coefficients a_k. Also, let

$$g[n] = x[n] - x[n-1].$$

(a) Show that $g[n]$ has a fundamental period of 10.
(b) Determine the Fourier series coefficients of $g[n]$.
(c) Using the Fourier series coefficients of $g[n]$ and the First-Difference property in Table 3.2, determine a_k for $k \neq 0$.

3.32. Consider the signal $x[n]$ depicted in Figure P3.32. This signal is periodic with period $N = 4$. The signal can be expressed in terms of a discrete-time Fourier series as

$$x[n] = \sum_{k=0}^{3} a_k e^{jk(2\pi/4)n}. \qquad \text{(P3.32–1)}$$

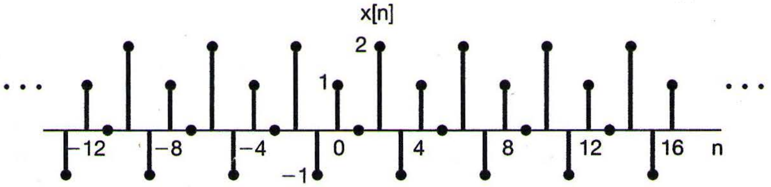


Figure P3.32

As mentioned in the text, one way to determine the Fourier series coefficients is to treat eq. (P3.32–1) as a set of four linear equations (for $n = 0, 1, 2, 3$) in four unknowns (a_0, a_1, a_2, and a_3).

(a) Write out these four equations explicitly, and solve them directly using any standard technique for solving four equations in four unknowns. (Be sure first to reduce the foregoing complex exponentials to the simplest form.)
(b) Check your answer by calculating the a_k directly, using the discrete-time Fourier series analysis equation

$$a_k = \frac{1}{4}\sum_{n=0}^{3} x[n]e^{-jk(2\pi/4)n}.$$

3.33. Consider a causal continuous-time LTI system whose input $x(t)$ and output $y(t)$ are related by the following differential equation:

$$\frac{d}{dt}y(t) + 4y(t) = x(t).$$

Find the Fourier series representation of the output $y(t)$ for each of the following inputs:

(a) $x(t) = \cos 2\pi t$

(b) $x(t) = \sin 4\pi t + \cos(6\pi t + \pi/4)$

3.34. Consider a continuous-time LTI system with impulse response

$$h(t) = e^{-4|t|}.$$

Find the Fourier series representation of the output $y(t)$ for each of the following inputs:

(a) $x(t) = \sum_{n=-\infty}^{+\infty}\delta(t-n)$

(b) $x(t) = \sum_{n=-\infty}^{+\infty}(-1)^n\delta(t-n)$

(c) $x(t)$ is the periodic wave depicted in Figure P3.34.

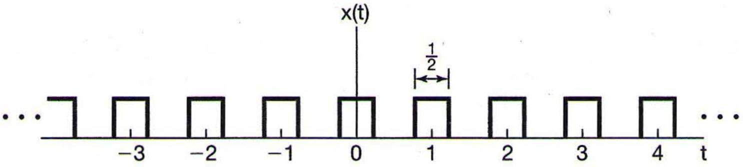


Figure P3.34

3.35. Consider a continuous-time LTI system S whose frequency response is

$$H(j\omega) = \begin{cases} 1, & |\omega| \geq 250 \\ 0, & \text{otherwise} \end{cases}.$$

When the input to this system is a signal $x(t)$ with fundamental period $T = \pi/7$ and Fourier series coefficients a_k, it is found that the output $y(t)$ is identical to $x(t)$. For what values of k is it guaranteed that $a_k = 0$?

3.36. Consider a causal discrete-time LTI system whose input $x[n]$ and output $y[n]$ are related by the following difference equation:

$$y[n] - \frac{1}{4}y[n-1] = x[n]$$

Find the Fourier series representation of the output $y[n]$ for each of the following inputs:

(a) $x[n] = \sin(\frac{3\pi}{4}n)$

(b) $x[n] = \cos(\frac{\pi}{4}n) + 2\cos(\frac{\pi}{2}n)$

3.37. Consider a discrete-time LTI system with impulse response

$$h[n] = \left(\frac{1}{2}\right)^{|n|}.$$

Find the Fourier series representation of the output $y[n]$ for each of the following inputs:

(a) $x[n] = \sum_{k=-\infty}^{\infty}\delta[n-4k]$

(b) $x[n]$ is periodic with period 6 and

$$x[n] = \begin{cases} 1, & n = 0, \pm 1 \\ 0, & n = \pm 2, \pm 3 \end{cases}$$

3.38. Consider a discrete-time LTI system with impulse response

$$h[n] = \begin{cases} 1, & 0 \le n \le 2 \\ -1, & -2 \le n \le -1 \\ 0, & \text{otherwise} \end{cases}.$$

Given that the input to this system is

$$x[n] = \sum_{k=-\infty}^{+\infty} \delta[n-4k],$$

determine the Fourier series coefficients of the output $y[n]$.

3.39. Consider a discrete-time LTI system S whose frequency response is

$$H(e^{j\omega}) = \begin{cases} 1, & |\omega| \le \frac{\pi}{8} \\ 0, & \frac{\pi}{8} < |\omega| < \pi \end{cases}.$$

Show that if the input $x[n]$ to this system has a period $N = 3$, the output $y[n]$ has only one nonzero Fourier series coefficient per period.

ADVANCED PROBLEMS

3.40. Let $x(t)$ be a periodic signal with fundamental period T and Fourier series coefficients a_k. Derive the Fourier series coefficients of each of the following signals in terms of a_k:

(a) $x(t-t_0) + x(t+t_0)$

(b) $\mathcal{E}v\{x(t)\}$

(c) $\mathcal{R}e\{x(t)\}$

(d) $\frac{d^2x(t)}{dt^2}$

(e) $x(3t-1)$ [for this part, first determine the period of $x(3t-1)$]

3.41. Suppose we are given the following information about a continuous-time periodic signal with period 3 and Fourier coefficients a_k:

1. $a_k = a_{k+2}$.
2. $a_k = a_{-k}$.
3. $\int_{-0.5}^{0.5} x(t)dt = 1$.
4. $\int_{1}^{2} x(t)dt = 2$.

Determine $x(t)$.

3.42. Let $x(t)$ be a real-valued signal with fundamental period T and Fourier series coefficients a_k.

(a) Show that $a_k = a^*_{-k}$ and a_0 must be real.

(b) Show that if $x(t)$ is even, then its Fourier series coefficients must be real and even.

(c) Show that if $x(t)$ is odd, then its Fourier series coefficients are imaginary and odd and $a_0 = 0$.

(d) Show that the Fourier coefficients of the even part of $x(t)$ are equal to $\mathcal{R}e\{a_k\}$.

(e) Show that the Fourier coefficients of the odd part of $x(t)$ are equal to $j\mathcal{I}m\{a_k\}$.

3.43. **(a)** A continuous-time periodic signal $x(t)$ with period T is said to be *odd harmonic* if, in its Fourier series representation

$$x(t) = \sum_{k=-\infty}^{+\infty} a_k e^{jk(2\pi/T)t}, \tag{P3.43–1}$$

$a_k = 0$ for every non-zero even integer k.

(i) Show that if $x(t)$ is odd harmonic, then

$$x(t) = -x\left(t + \frac{T}{2}\right). \tag{P3.43–2}$$

(ii) Show that if $x(t)$ satisfies eq. (P3.43–2), then it is odd harmonic.

(b) Suppose that $x(t)$ is an odd-harmonic periodic signal with period 2 such that

$$x(t) = t \quad \text{for } 0 < t < 1.$$

Sketch $x(t)$ and find its Fourier series coefficients.

(c) Analogously, to an odd-harmonic signal, we could define an even-harmonic signal as a signal for which $a_k = 0$ for k odd in the representation in eq. (P3.43–1). Could T be the fundamental period for such a signal? Explain your answer.

(d) More generally, show that T is the fundamental period of $x(t)$ in eq. (P3.43–1) if one of two things happens:

(1) Either a_1 or a_{-1} is nonzero;

or

(2) There are two integers k and l that have no common factors and are such that both a_k and a_l are nonzero.

3.44. Suppose we are given the following information about a signal $x(t)$:

1. $x(t)$ is a real signal.
2. $x(t)$ is periodic with period $T = 6$ and has Fourier coefficients a_k.
3. $a_k = 0$ for $k = 0$ and $k > 2$.
4. $x(t) = -x(t-3)$.
5. $\frac{1}{6}\int_{-3}^{3}|x(t)|^2\,dt = \frac{1}{2}$.
6. a_1 is a positive real number.

Show that $x(t) = A\cos(Bt + C)$, and determine the values of the constants A, B, and C.

3.45. Let $x(t)$ be a real periodic signal with Fourier series representation given in the sine-cosine form of eq. (3.32); i.e.,

$$x(t) = a_0 + 2\sum_{k=1}^{\infty}[B_k \cos k\omega_0 t - C_k \sin k\omega_0 t]. \qquad \text{(P3.45–1)}$$

(a) Find the exponential Fourier series representation of the even and odd parts of $x(t)$; that is, find the coefficients α_k and β_k in terms of the coefficients in eq. (P3.45–1) so that

$$\mathcal{E}v\{x(t)\} = \sum_{k=-\infty}^{+\infty} \alpha_k e^{jk\omega_0 t},$$

$$\mathcal{O}d\{x(t)\} = \sum_{k=-\infty}^{+\infty} \beta_k e^{jk\omega_0 t}.$$

(b) What is the relationship between α_k and α_{-k} in part (a)? What is the relationship between β_k and β_{-k}?

(c) Suppose that the signals $x(t)$ and $z(t)$ shown in Figure P3.45 have the sine-cosine series representations

$$x(t) = a_0 + 2\sum_{k=1}^{\infty}\left[B_k \cos\left(\frac{2\pi kt}{3}\right) - C_k \sin\left(\frac{2\pi kt}{3}\right)\right],$$

$$z(t) = d_0 + 2\sum_{k=1}^{\infty}\left[E_k \cos\left(\frac{2\pi kt}{3}\right) - F_k \sin\left(\frac{2\pi kt}{3}\right)\right].$$

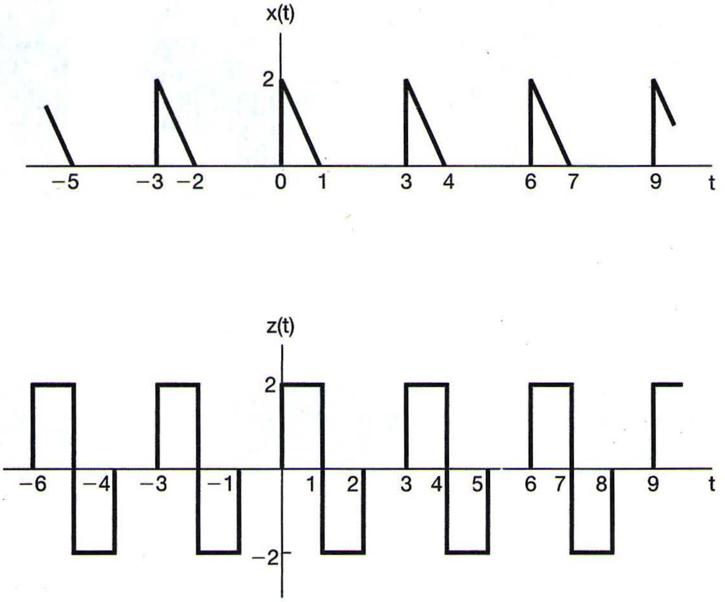


Figure P3.45

Sketch the signal

$$y(t) = 4(a_0 + d_0) + 2\sum_{k=1}^{\infty}\left\{\left[B_k + \frac{1}{2}E_k\right]\cos\left(\frac{2\pi kt}{3}\right) + F_k \sin\left(\frac{2\pi kt}{3}\right)\right\}.$$

3.46 In this problem, we derive two important properties of the continuous-time Fourier series: the multiplication property and Parseval's relation. Let $x(t)$ and $y(t)$ both be continuous-time periodic signals having period T_0 and with Fourier series representations given by

$$x(t) = \sum_{k=-\infty}^{+\infty} a_k e^{jk\omega_0 t}, \quad y(t) = \sum_{k=-\infty}^{+\infty} b_k e^{jk\omega_0 t}. \tag{P3.46–1}$$

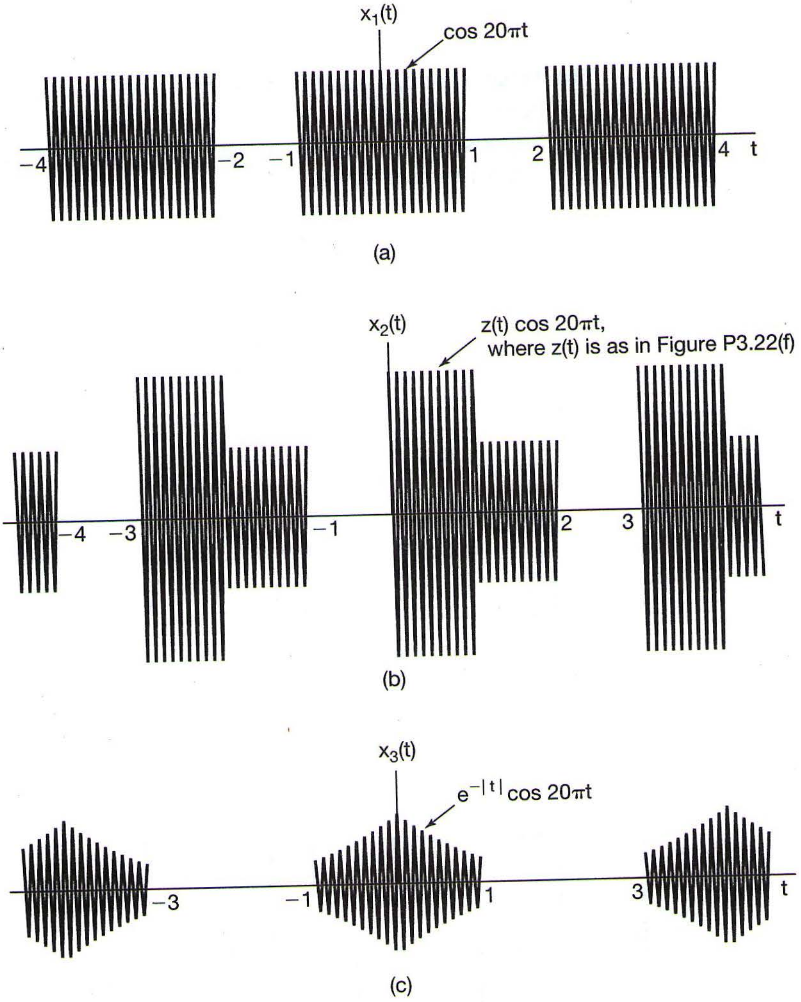


Figure P3.46

(a) Show that the Fourier series coefficients of the signal

$$z(t) = x(t)y(t) = \sum_{k=-\infty}^{+\infty} c_k e^{jk\omega_0 t}$$

are given by the discrete convolution

$$c_k = \sum_{n=-\infty}^{+\infty} a_n b_{k-n}.$$

(b) Use the result of part (a) to compute the Fourier series coefficients of the signals $x_1(t)$, $x_2(t)$, and $x_3(t)$ depicted in Figure P3.46.

(c) Suppose that $y(t)$ in eq. (P3.46–1) equals $x^*(t)$. Express the b_k in the equation in terms of a_k, and use the result of part (a) to prove Parseval's relation for periodic signals—that is,

$$\frac{1}{T_0}\int_0^{T_0} |x(t)|^2\, dt = \sum_{k=-\infty}^{+\infty} |a_k|^2.$$

3.47 Consider the signal

$$x(t) = \cos 2\pi t.$$

Since $x(t)$ is periodic with a fundamental period of 1, it is also periodic with a period of N, where N is any positive integer. What are the Fourier series coefficients of $x(t)$ if we regard it as a periodic signal with period 3?

3.48. Let $x[n]$ be a periodic sequence with period N and Fourier series representation

$$x[n] = \sum_{k=<N>} a_k e^{jk(2\pi/N)n}. \qquad \text{(P3.48–1)}$$

The Fourier series coefficients for each of the following signals can be expressed in terms of a_k in eq. (P3.48–1). Derive the expressions.

(a) $x[n - n_0]$

(b) $x[n] - x[n-1]$

(c) $x[n] - x[n - \frac{N}{2}]$ (assume that N is even)

(d) $x[n] + x[n + \frac{N}{2}]$ (assume that N is even; note that this signal is periodic with period $N/2$)

(e) $x^*[-n]$

(f) $(-1)^n x[n]$ (assume that N is even)

(g) $(-1)^n x[n]$ (assume that N is odd; note that this signal is periodic with period $2N$)

(h) $y[n] = \begin{cases} x[n], & n \text{ even} \\ 0, & n \text{ odd} \end{cases}$

3.49. Let $x[n]$ be a periodic sequence with period N and Fourier series representation

$$x[n] = \sum_{k=<N>} a_k e^{jk(2\pi/N)n}. \qquad \text{(P3.49–1)}$$

(a) Suppose that N is even and that $x[n]$ in eq. (P3.49–1) satisfies

$$x[n] = -x\left[n + \frac{N}{2}\right] \text{ for all } n.$$

Show that $a_k = 0$ for all even integers k.

(b) Suppose that N is divisible by 4. Show that if

$$x[n] = -x\left[n + \frac{N}{4}\right] \text{ for all } n,$$

then $a_k = 0$ for every value of k that is a multiple of 4.

(c) More generally, suppose that N is divisible by an integer M. Show that if

$$\sum_{r=0}^{(N/M)-1} x\left[n + r\frac{N}{M}\right] = 0 \text{ for all } n,$$

then $a_k = 0$ for every value of k that is a multiple of M.

3.50. Suppose we are given the following information about a periodic signal $x[n]$ with period 8 and Fourier coefficients a_k:

1. $a_k = -a_{k-4}$.
2. $x[2n + 1] = (-1)^n$.

Sketch one period of $x[n]$.

3.51. Let $x[n]$ be a periodic signal with period $N = 8$ and Fourier series coefficients $a_k = -a_{k-4}$. A signal

$$y[n] = \left(\frac{1 + (-1)^n}{2}\right)x[n - 1]$$

with period $N = 8$ is generated. Denoting the Fourier series coefficients of $y[n]$ by b_k, find a function $f[k]$ such that

$$b_k = f[k]a_k.$$

3.52. $x[n]$ is a real periodic signal with period N and complex Fourier series coefficients a_k. Let the Cartesian form for a_k be denoted by

$$a_k = b_k + jc_k,$$

where b_k and c_k are both real.

(a) Show that $a_{-k} = a_k^*$. What is the relation between b_k and b_{-k}? What is the relation between c_k and c_{-k}?

(b) Suppose that N is even. Show that $a_{N/2}$ is real.

(c) Show that $x[n]$ can also be expressed as a trigonometric Fourier series of the form

$$x[n] = a_0 + 2\sum_{k=1}^{(N-1)/2} b_k \cos\left(\frac{2\pi kn}{N}\right) - c_k \sin\left(\frac{2\pi kn}{N}\right)$$

if N is odd or as

$$x[n] = (a_0 + a_{N/2}(-1)^n) + 2\sum_{k=1}^{(N-2)/2} b_k \cos\left(\frac{2\pi kn}{N}\right) - c_k \sin\left(\frac{2\pi kn}{N}\right)$$

if N is even.

(d) Show that if the polar form of a_k is $A_k e^{j\theta k}$, then the Fourier series representation for $x[n]$ can also be written as

$$x[n] = a_0 + 2\sum_{k=1}^{(N-1)/2} A_k \cos\left(\frac{2\pi kn}{N} + \theta_k\right)$$

if N is odd or as

$$x[n] = (a_0 + a_{N/2}(-1)^n) + 2\sum_{k=1}^{(N/2)-1} A_k \cos\left(\frac{2\pi kn}{N} + \theta_k\right)$$

if N is even.

(e) Suppose that $x[n]$ and $z[n]$, as depicted in Figure P3.52, have the sine-cosine series representations

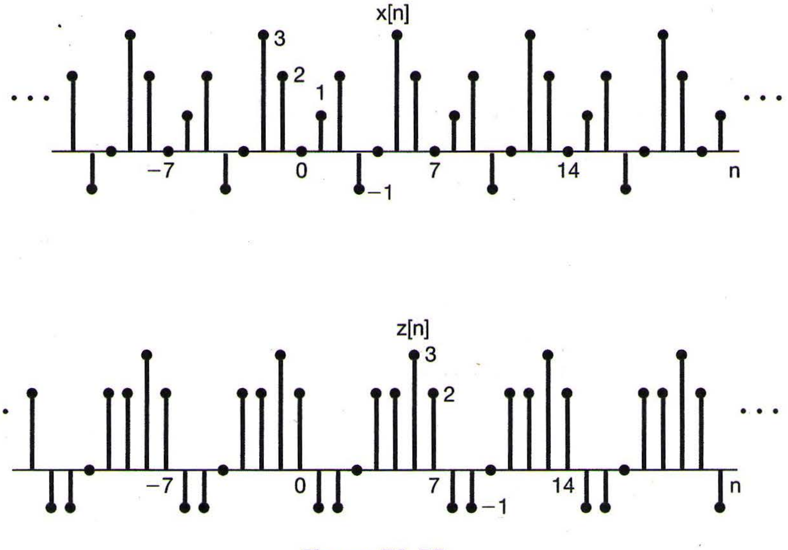


Figure P3.52

$$x[n] = a_0 + 2\sum_{k=1}^{3}\left\{ b_k \cos\left(\frac{2\pi kn}{7}\right) - c_k \sin\left(\frac{2\pi kn}{7}\right)\right\},$$

$$z[n] = d_0 + 2\sum_{k=1}^{3}\left\{ d_k \cos\left(\frac{2\pi kn}{7}\right) - f_k \sin\left(\frac{2\pi kn}{7}\right)\right\}.$$

Sketch the signal

$$y[n] = a_0 - d_0 + 2\sum_{k=1}^{3}\left\{ d_k \cos\left(\frac{2\pi kn}{7}\right) + (f_k - c_k)\sin\left(\frac{2\pi kn}{7}\right)\right\}.$$

3.53. Let $x[n]$ be a real periodic signal with period N and Fourier coefficients a_k.

(a) Show that if N is even, at least two of the Fourier coefficients within one period of a_k are real.

(b) Show that if N is odd, at least one of the Fourier coefficients within one period of a_k is real.

3.54. Consider the function

$$a[k] = \sum_{n=0}^{N-1} e^{j(2\pi/N)kn}.$$

(a) Show that $a[k] = N$ for $k = 0, \pm N, \pm 2N, \pm 3N, \ldots$.

(b) Show that $a[k] = 0$ whenever k is not an integer multiple of N. (*Hint:* Use the finite sum formula.)

(c) Repeat parts (a) and (b) if

$$a[k] = \sum_{n=<N>} e^{j(2\pi/N)kn}.$$

3.55. Let $x[n]$ be a periodic signal with fundamental period N and Fourier series coefficients a_k. In this problem, we derive the time-scaling property

$$x_{(m)}[n] = \begin{cases} x[\frac{n}{m}], & n = 0, \pm m, \pm 2m, \cdots \\ 0, & \text{elsewhere} \end{cases}$$

listed in Table 3.2.

(a) Show that $x_{(m)}[n]$ has period of mN.

(b) Show that if

$$x[n] = v[n] + w[n],$$

then

$$x_{(m)}[n] = v_{(m)}[n] + w_{(m)}[n].$$

(c) Assuming that $x[n] = e^{j2\pi k_0 n/N}$ for some integer k_0, verify that

$$x_{(m)}[n] = \frac{1}{m}\sum_{l=0}^{m-1} e^{j2\pi(k_0+lN)n/(mN)}.$$

That is, one complex exponential in $x[n]$ becomes a linear combination of m complex exponentials in $x_{(m)}[n]$.

(d) Using the results of parts (a), (b), and (c), show that if $x[n]$ has the Fourier coefficients a_k, then $x_{(m)}[n]$ must have the Fourier coefficients $\frac{1}{m}a_k$.

3.56. Let $x[n]$ be a periodic signal with period N and Fourier coefficients a_k.

(a) Express the Fourier coefficients b_k of $|x[n]|^2$ in terms of a_k.

(b) If the coefficients a_k are real, is it guaranteed that the coefficients b_k are also real?

3.57. **(a)** Let

$$x[n] = \sum_{k=0}^{N-1} a_k e^{jk(2\pi/N)n} \tag{P3.57–1}$$

and

$$y[n] = \sum_{k=0}^{N-1} b_k e^{jk(2\pi/N)n}$$

be periodic signals. Show that

$$x[n]y[n] = \sum_{k=0}^{N-1} c_k e^{jk(2\pi/N)n},$$

where

$$c_k = \sum_{l=0}^{N-1} a_l b_{k-l} = \sum_{l=0}^{N-1} a_{k-l} b_l.$$

(b) Generalize the result of part (a) by showing that

$$c_k = \sum_{l=<N>} a_l b_{k-l} = \sum_{l=<N>} a_{k-l} b_l.$$

(c) Use the result of part (b) to find the Fourier series representation of the following signals, where $x[n]$ is given by eq. (P3.57–1).

(i) $x[n]\cos\left(\frac{6\pi n}{N}\right)$

(ii) $x[n]\sum_{r=-\infty}^{+\infty}\delta[n - rN]$

(iii) $x[n]\left(\sum_{r=-\infty}^{+\infty}\delta\left[n - \frac{rN}{3}\right]\right)$ (assume that N is divisible by 3)

(d) Find the Fourier series representation for the signal $x[n]y[n]$, where

$$x[n] = \cos(\pi n/3)$$

and

$$y[n] = \begin{cases} 1, & |n| \leq 3 \\ 0, & 4 \leq |n| \leq 6 \end{cases}$$

is periodic with period 12.

(e) Use the result of part (b) to show that

$$\sum_{n=<N>} x[n]y[n] = N \sum_{l=<N>} a_l b_{-l},$$

and from this expression, derive Parseval's relation for discrete-time periodic signals.

3.58. Let $x[n]$ and $y[n]$ be periodic signals with common period N, and let

$$z[n] = \sum_{r=<N>} x[r]y[n-r]$$

be their periodic convolution.

(a) Show that $z[n]$ is also periodic with period N.

(b) Verify that if a_k, b_k, and c_k are the Fourier coefficients of $x[n]$, $y[n]$, and $z[n]$, respectively, then

$$c_k = N a_k b_k.$$

(c) Let

$$x[n] = \sin\left(\frac{3\pi n}{4}\right)$$

and

$$y[n] = \begin{cases} 1, & 0 \leq n \leq 3 \\ 0, & 4 \leq n \leq 7 \end{cases}$$

be two signals that are periodic with period 8. Find the Fourier series representation for the periodic convolution of these signals.

(d) Repeat part (c) for the following two periodic signals that also have period 8:

$$x[n] = \begin{cases} \sin\left(\frac{3\pi n}{4}\right), & 0 \leq n \leq 3 \\ 0, & 4 \leq n \leq 7 \end{cases},$$

$$y[n] = \left(\frac{1}{2}\right)^n, 0 \leq n \leq 7.$$

3.59. **(a)** Suppose $x[n]$ is a periodic signal with period N. Show that the Fourier series coefficients of the periodic signal

$$g(t) = \sum_{k=-\infty}^{\infty} x[k]\,\delta(t-kT)$$

are periodic with period N.

(b) Suppose that $x(t)$ is a periodic signal with period T and Fourier series coefficients a_k with period N. Show that there must exist a periodic sequence $g[n]$ such that

$$x(t) = \sum_{k=-\infty}^{\infty} g[k]\,\delta(t - kT/N).$$

(c) Can a continuous periodic signal have periodic Fourier coefficients?

3.60. Consider the following pairs of signals $x[n]$ and $y[n]$. For each pair, determine whether there is a discrete-time LTI system for which $y[n]$ is the output when the corresponding $x[n]$ is the input. If such a system exists, determine whether the system is unique (i.e., whether there is more than one LTI system with the given input-output pair). Also, determine the frequency response of an LTI system with the desired behavior. If no such LTI system exists for a given $x[n]$, $y[n]$ pair, explain why.

(a) $x[n] = (\frac{1}{2}^n)$, $y[n] = (\frac{1}{4}^n)$
(b) $x[n] = (\frac{1}{2}^n)u[n]$, $y[n] = (\frac{1}{4})^n u[n]$
(c) $x[n] = (\frac{1}{2}^n)u[n]$, $y[n] = 4^n u[-n]$
(d) $x[n] = e^{jn/8}$, $y[n] = 2e^{jn/8}$
(e) $x[n] = e^{jn/8}u[n]$, $y[n] = 2e^{jn/8}u[n]$
(f) $x[n] = j^n$, $y[n] = 2j^n(1 - j)$
(g) $x[n] = \cos(\pi n/3), y[n] = \cos(\pi n/3) + \sqrt{3}\sin(\pi n/3)$
(h) $x[n]$ and $y_1[n]$ as in Figure P3.60
(i) $x[n]$ and $y_2[n]$ as in Figure P3.60

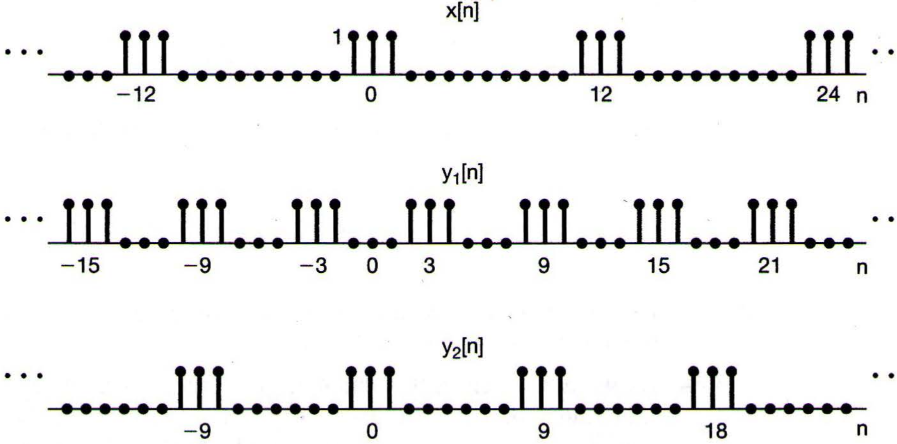


Figure P3.60

3.61. As we have seen, the techniques of Fourier analysis are of value in examining continuous-time LTI systems because periodic complex exponentials are eigenfunctions for LTI systems. In this problem, we wish to substantiate the following statement: Although some LTI systems may have additional eigenfunctions, the complex exponentials are the *only* signals that are eigenfunctions of *every* LTI system.

(a) What are the eigenfunctions of the LTI system with unit impulse response $h(t) = \delta(t)$? What are the associated eigenvalues?

(b) Consider the LTI system with unit impulse response $h(t) = \delta(t - T)$. Find a signal that is not of the form e^{st}, but that is an eigenfunction of the system with eigenvalue 1. Similarly, find the eigenfunctions with eigenvalues 1/2 and 2 that are not complex exponentials. (*Hint:* You can find impulse trains that meet these requirements.)

(c) Consider a stable LTI system with impulse response $h(t)$ that is real and even. Show that $\cos \omega t$ and $\sin \omega t$ are eigenfunctions of this system.

(d) Consider the LTI system with impulse response $h(t) = u(t)$. Suppose that $\phi(t)$ is an eigenfunction of this system with eigenvalue λ. Find the differential equation that $\phi(t)$ must satisfy, and solve the equation. This result, together with those of parts (a) through (c), should prove the validity of the statement made at the beginning of the problem.

3.62. One technique for building a dc power supply is to take an ac signal and full-wave rectify it. That is, we put the ac signal $x(t)$ through a system that produces $y(t) = |x(t)|$ as its output.

(a) Sketch the input and output waveforms if $x(t) = \cos t$. What are the fundamental periods of the input and output?

(b) If $x(t) = \cos t$, determine the coefficients of the Fourier series for the output $y(t)$.

(c) What is the amplitude of the dc component of the input signal? What is the amplitude of the dc component of the output signal?

3.63. Suppose that a continuous-time periodic signal is the input to an LTI system. The signal has a Fourier series representation

$$x(t) = \sum_{k=-\infty}^{\infty} \alpha^{|k|} e^{jk(\pi/4)t},$$

where α is a real number between 0 and 1, and the frequency response of the system is

$$H(j\omega) = \begin{cases} 1, & |\omega| \leq W \\ 0, & |\omega| > W \end{cases}.$$

How large must W be in order for the output of the system to have at least 90% of the average energy per period of $x(t)$?

3.64. As we have seen in this chapter, the concept of an eigenfunction is an extremely important tool in the study of LTI systems. The same can be said for linear, but time-varying, systems. Specifically, consider such a system with input $x(t)$ and output $y(t)$. We say that a signal $\phi(t)$ is an *eigenfunction* of the system if

$$\phi(t) \longrightarrow \lambda\phi(t).$$

That is, if $x(t) = \phi(t)$, then $y(t) = \lambda\phi(t)$, where the complex constant λ is called the *eigenvalue associated with* $\phi(t)$.

(a) Suppose that we can represent the input $x(t)$ to our system as a linear combination of eigenfunctions $\phi_k(t)$, each of which has a corresponding eigenvalue λ_k; that is,

$$x(t) = \sum_{k=-\infty}^{\infty} c_k \phi_k(t).$$

Express the output $y(t)$ of the system in terms of $\{c_k\}$, $\{\phi_k(t)\}$, and $\{\lambda_k\}$.

(b) Consider the system characterized by the differential equation

$$y(t) = t^2 \frac{d^2 x(t)}{dt^2} + t\frac{dx(t)}{dt}.$$

Is this system linear? Is it time invariant?

(c) Show that the functions

$$\phi_k(t) = t^k$$

are eigenfunctions of the system in part (b). For each $\phi_k(t)$, determine the corresponding eigenvalue λ_k.

(d) Determine the output of the system if

$$x(t) = 10t^{-10} + 3t + \frac{1}{2}t^4 + \pi.$$

EXTENSION PROBLEMS

3.65. Two functions $u(t)$ and $v(t)$ are said to be *orthogonal over the interval (a,b)* if

$$\int_a^b u(t)v^*(t)\,dt = 0. \qquad \text{(P3.65–1)}$$

If, in addition,

$$\int_a^b |u(t)|^2\,dt = 1 = \int_a^b |v(t)|^2\,dt,$$

the functions are said to be *normalized* and hence are called *orthonormal.* A set of functions $\{\phi_k(t)\}$ is called an *orthogonal (orthonormal) set* if each pair of functions in the set is orthogonal (orthonormal).

(a) Consider the pairs of signals $u(t)$ and $v(t)$ depicted in Figure P3.65. Determine whether each pair is orthogonal over the interval (0, 4).

(b) Are the functions $\sin m\omega_0 t$ and $\sin n\omega_0 t$ orthogonal over the interval $(0, T)$, where $T = 2\pi/\omega_0$? Are they also orthonormal?

(c) Repeat part (b) for the functions $\phi_m(t)$ and $\phi_n(t)$, where

$$\phi_k(t) = \frac{1}{\sqrt{T}}[\cos k\omega_0 t + \sin k\omega_0 t].$$

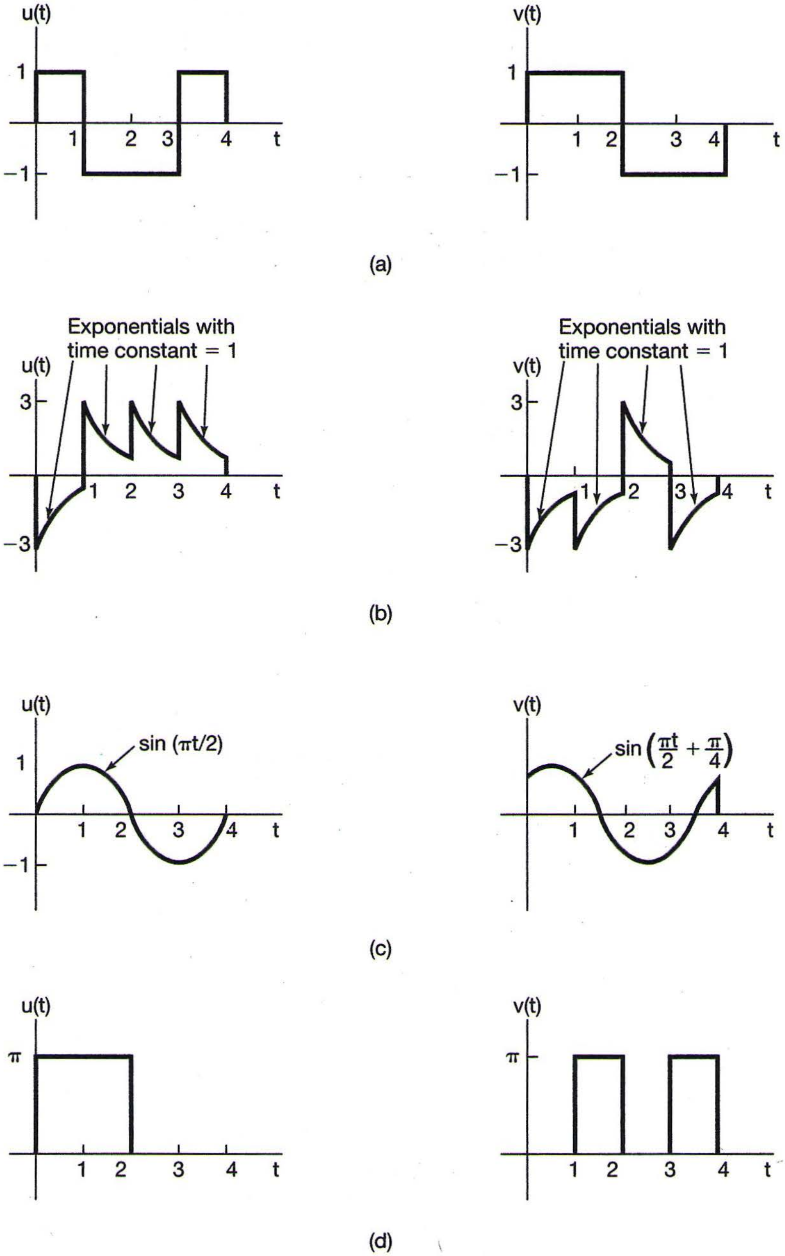


Figure P3.65

(d) Show that the functions $\phi_k(t) = e^{jk\omega_0 t}$ are orthogonal over *any* interval of length $T = 2\pi/\omega_0$. Are they orthonormal?

(e) Let $x(t)$ be an arbitrary signal, and let $x_o(t)$ and $x_e(t)$ be, respectively, the odd and even parts of $x(t)$. Show that $x_o(t)$ and $x_e(t)$ are orthogonal over the interval $(-T, T)$ for any T.

(f) Show that if $\{\phi_k(t)\}$ is a set of orthogonal signals over the interval (a, b), then the set $\{(1/\sqrt{A_k})\phi_k(t)\}$, where

$$A_k = \int_a^b |\phi_k(t)|^2\,dt,$$

is orthonormal.

(g) Let $\{\phi_i(t)\}$ be a set of orthonormal signals on the interval (a, b), and consider a signal of the form

$$x(t) = \sum_i a_i\phi_i(t),$$

where the a_i are complex constants. Show that

$$\int_a^b |x(t)|^2\,dt = \sum_i |a_i|^2.$$

(h) Suppose that $\phi_1(t), \ldots, \phi_N(t)$ are nonzero only in the time interval $0 \le t \le T$ and that they are orthonormal over this time interval. Let L_i denote the LTI system with impulse response

$$h_i(t) = \phi_i(T - t). \qquad \text{(P3.65–2)}$$

Show that if $\phi_j(t)$ is applied to this system, then the output at time T is 1 if $i = j$ and 0 if $i \neq j$. The system with impulse response given by eq. (P3.65–2) was referred to in Problems 2.66 and 2.67 as the *matched filter* for the signal $\phi_i(t)$.

3.66. The purpose of this problem is to show that the representation of an arbitrary periodic signal by a Fourier series or, more generally, as a linear combination of any set of orthogonal functions is computationally efficient and in fact very useful for obtaining good approximations of signals.[12]

Specifically, let $\{\phi_i(t)\}$, $i = 0, \pm 1, \pm 2, \ldots$ be a set of orthonormal functions on the interval $a \le t \le b$, and let $x(t)$ be a given signal. Consider the following approximation of $x(t)$ over the interval $a \le t \le b$:

$$\hat{x}_n(t) = \sum_{i=-N}^{+N} a_i\phi_i(t). \qquad \text{(P3.66–1)}$$

Here, the a_i are (in general, complex) constants. To measure the deviation between $x(t)$ and the series approximation $\hat{x}_N(t)$, we consider the error $e_N(t)$ defined as

$$e_N(t) = x(t) - \hat{x}_N(t). \qquad \text{(P3.66–2)}$$

A reasonable and widely used criterion for measuring the quality of the approximation is the energy in the error signal over the interval of interest—that is, the integral of the square of the magnitude of the error over the interval $a \le t \le b$:

---

[12] See Problem 3.65 for the definitions of orthogonal and orthonormal functions.

$$E = \int_a^b |e_N(t)|^2\, dt. \tag{P3.66–3}$$

(a) Show that E is minimized by choosing

$$a_i = \int_a^b x(t)\phi_i^*(t)\, dt. \tag{P3.66–4}$$

[*Hint:* Use eqs. (P3.66–1)–(P3.66–3) to express E in terms of a_i, $\phi_i(t)$, and $x(t)$. Then express a_i in rectangular coordinates as $a_i = b_i + jc_i$, and show that the equations

$$\frac{\partial E}{\partial b_i} = 0 \quad \text{and} \quad \frac{\partial E}{\partial c_i} = 0,\ i = 0, \pm 1, \pm 2, \ldots, N$$

are satisfied by the a_i as given by eq. (P3.66–4).]

(b) How does the result of part (a) change if

$$A_i = \int_a^b |\phi_i(t)|^2\, dt$$

and the $\{\phi_i(t)\}$ are orthogonal but not orthonormal?

(c) Let $\phi_n(t) = e^{jn\omega_0 t}$, and choose any interval of length $T_0 = 2\pi/\omega_0$. Show that the a_i that minimize E are as given in eq. (3.50).

(d) The set of *Walsh functions* is an often-used set of orthonormal functions. (See Problem 2.66.) The set of five Walsh functions, $\phi_0(t)$, $\phi_1(t)$, ..., $\phi_4(t)$, is illustrated in Figure P3.66, where we have scaled time so that the $\phi_i(t)$ are nonzero and orthonormal over the interval $0 \le t \le 1$. Let $x(t) = \sin \pi t$. Find the approximation of $x(t)$ of the form

$$\hat{x}(t) = \sum_{i=0}^{4} a_i \phi_i(t)$$

such that

$$\int_0^1 |x(t) - \hat{x}(t)|^2\, dt$$

is minimized.

(e) Show that $\hat{x}_N(t)$ in eq. (P3.66–1) and $e_N(t)$ in eq. (P3.66–2) are orthogonal if the a_i are chosen as in eq. (P3.66–4).

The results of parts (a) and (b) are extremely important in that they show that each coefficient a_i is *independent* of all the other a_j's, $i \neq j$. Thus, if we add more terms to the approximation [e.g., if we compute the approximation $\hat{x}_{N+1}(t)$], the coefficients of $\phi_i(t)$, $i = 1, \ldots, N$, that were previously determined will not change. In contrast to this, consider another type of se-

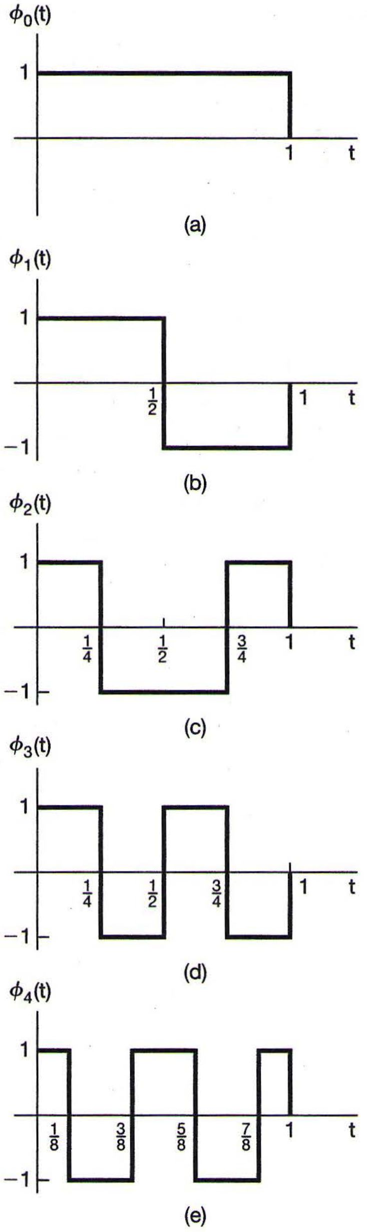


Figure P3.66

ries expansion, the polynomial Taylor series. The *infinite* Taylor series for e^t is $e^t = 1 + t + t^2/2! + \ldots$, but as we shall show, when we consider a *finite* polynomial series and the error criterion of eq. (P3.66–3), we get a very different result.

Specifically, let $\phi_0(t) = 1$, $\phi_1(t) = t$, $\phi_2(t) = t^2$, and so on.

(f) Are the $\phi_i(t)$ orthogonal over the interval $0 \leq t \leq 1$?

(g) Consider an approximation of $x(t) = e^t$ over the interval $0 \leq t \leq 1$ of the form

$$\hat{x}_0(t) = a_0\phi_0(t).$$

Find the value of a_0 that minimizes the energy in the error signal over the interval.

(h) We now wish to approximate e^t by a Taylor series using two terms—i.e., $\hat{x}_1(t) = a_0 + a_1 t$. Find the optimum values for a_0 and a_1. [*Hint:* Compute E in terms of a_0 and a_1, and then solve the simultaneous equations

$$\frac{\partial E}{\partial a_0} = 0 \quad \text{and} \quad \frac{\partial E}{\partial a_1} = 0.$$

Note that your answer for a_0 has changed from its value in part (g), where there was only one term in the series. Further, as you increase the number of terms in the series, that coefficient and all others will continue to change. We can thus see the advantage to be gained in expanding a function using orthogonal terms.]

3.67 As we discussed in the text, the origins of Fourier analysis can be found in problems of mathematical physics. In particular, the work of Fourier was motivated by his investigation of heat diffusion. In this problem, we illustrate how the Fourier series enter into the investigation.[13]

Consider the problem of determining the temperature at a given depth beneath the surface of the earth as a function of time, where we assume that the temperature at the surface is a given function of time $T(t)$ that is periodic with period 1. (The unit of time is one year.) Let $T(x, t)$ denote the temperature at a depth x below the surface at time t. This function obeys the heat diffusion equation

$$\frac{\partial T(x,t)}{\partial t} = \frac{1}{2}k^2\frac{\partial^2 T(x,t)}{\partial x^2} \tag{P3.67–1}$$

with auxiliary condition

$$T(0,t) = T(t). \tag{P3.67–2}$$

Here, k is the heat diffusion constant for the earth ($k > 0$). Suppose that we expand $T(t)$ in a Fourier series:

$$T(t) = \sum_{n=-\infty}^{+\infty} a_n e^{jn2\pi t}. \tag{P3.67–3}$$

Similarly, let us expand $T(x, t)$ at any given depth x in a Fourier series in t. We obtain

$$T(x,t) = \sum_{n=-\infty}^{+\infty} b_n(x) e^{jn2\pi t}, \tag{P3.67–4}$$

where the Fourier coefficients $b_n(x)$ depend upon the depth x.

---

[13] The problem has been adapted from A. Sommerfeld, *Partial Differential Equations in Physics* (New York: Academic Press, 1949), pp 68–71.

(a) Use eqs. (P3.67–1)–(P3.67–4) to show that $b_n(x)$ satisfies the differential equation

$$\frac{d^2 b_n(x)}{dx^2} = \frac{4\pi jn}{k^2} b_n(x) \tag{P3.67–5a}$$

with auxiliary condition

$$b_n(0) = a_n. \tag{P3.67–5b}$$

Since eq. (P3.67–5a) is a second-order equation, we need a second auxiliary condition. We argue on physical grounds that, far below the earth's surface, the variations in temperature due to surface fluctuations should disappear. That is,

$$\lim_{x\to\infty} T(x, t) = \text{a constant.} \tag{P3.67–5c}$$

(b) Show that the solution of eqs. (P3.67–5) is

$$b_n(x) = \begin{cases} a_n \exp[-\sqrt{2\pi|n|}(1 + j)x/k], & n \geq 0 \\ a_n \exp[-\sqrt{2\pi|n|}(1 - j)x/k], & n \leq 0 \end{cases}.$$

(c) Thus, the temperature oscillations at depth x are damped and phase-shifted versions of the temperature oscillations at the surface. To see this more clearly, let

$$T(t) = a_0 + a_1 \sin 2\pi t$$

(so that a_0 represents the mean yearly temperature). Sketch $T(t)$ and $T(x, t)$ over a one-year period for

$$x = k\sqrt{\frac{\pi}{2}},$$

$a_0 = 2$, and $a_1 = 1$. Note that at this depth not only are the temperature oscillations significantly damped, but the phase shift is such that it is warmest in winter and coldest in summer. This is exactly the reason why vegetable cellars are constructed!

3.68. Consider the closed contour shown in Figure P3.68. As illustrated, we can view this curve as being traced out by the tip of a rotating vector of varying length. Let $r(\theta)$ denote the length of the vector as a function of the angle θ. Then $r(\theta)$ is periodic in θ with period 2π and thus has a Fourier series representation. Let $\{a_k\}$ denote the Fourier coefficients of $r(\theta)$.

(a) Consider now the projection $x(\theta)$ of the vector $r(\theta)$ onto the x-axis, as indicated in the figure. Determine the Fourier coefficients for $x(\theta)$ in terms of the a_k's.

(b) Consider the sequence of coefficients

$$b_k = a_k e^{jk\pi/4}.$$

Sketch the figure in the plane that corresponds to this set of coefficients.

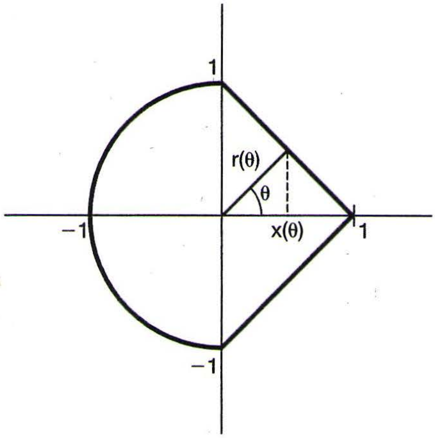


Figure P3.68

(c) Repeat part (b) with

$$b_k = a_k\delta[k].$$

(d) Sketch figures in the plane such that $r(\theta)$ is *not* constant, but does have each of the following properties:
(i) $r(\theta)$ is even.
(ii) The fundamental period of $r(\theta)$ is π.
(iii) The fundamental period of $r(\theta)$ is $\pi/2$.

3.69. In this problem, we consider the discrete-time counterpart of the concepts introduced in Problems 3.65 and 3.66. In analogy with the continuous-time case, two discrete-time signals $\phi_k[n]$ and $\phi_m[n]$ are said to be *orthogonal* over the interval (N_1, N_2) if

$$\sum_{n=N_1}^{N_2} \phi_k[n]\phi_m^*[n] = \begin{cases} A_k, & k = m \\ 0, & k \neq m \end{cases}. \tag{P3.69–1}$$

If the value of the constants A_k and A_m are both 1, then the signals are said to be *orthonormal.*

(a) Consider the signals

$$\phi_k[n] = \delta[n-k],\ k = 0, \pm 1, \pm 2, \ldots, \pm N.$$

Show that these signals are orthonormal over the interval $(-N, N)$.

(b) Show that the signals

$$\phi_k[n] = e^{jk(2\pi/N)n},\ k = 0, 1, \ldots, N-1,$$

are orthogonal over any interval of length N.

(c) Show that if

$$x[n] = \sum_{i=1}^{M} a_i\phi_i[n],$$

where the $\phi_i[n]$ are orthogonal over the interval (N_1, N_2), then

$$\sum_{n=N_1}^{N_2} |x[n]|^2 = \sum_{i=1}^{M} |a_i|^2 A_i.$$

(d) Let $\phi_i[n]$, $i = 0, 1, \ldots, M$, be a set of orthogonal functions over the interval (N_1, N_2), and let $x[n]$ be a given signal. Suppose that we wish to approximate $x[n]$ as a linear combination of the $\phi_i[n]$; that is,

$$\hat{x}[n] = \sum_{i=0}^{M} a_i \phi_i[n],$$

where the a_i are constant coefficients. Let

$$e[n] = x[n] - \hat{x}[n],$$

and show that if we wish to minimize

$$E = \sum_{n=N_1}^{N_2} |e[n]|^2,$$

then the a_i are given by

$$a_i = \frac{1}{A_i} \sum_{n=N_1}^{N_2} x[n]\phi_i^*[n]. \qquad \text{(P3.69–2)}$$

[*Hint:* As in Problem 3.66, express E in terms of a_i, $\phi_i[n]$, A_i, and $x[n]$, write $a_i = b_i + jc_i$, and show that the equations

$$\frac{\partial E}{\partial b_i} = 0 \quad \text{and} \quad \frac{\partial E}{\partial c_i} = 0$$

are satisfied by the a_i given by eq. (P3.69–2). Note that applying this result when the $\phi_i[n]$ are as in part (b) yields eq. (3.95) for a_k.]

(e) Apply the result of part (d) when the $\phi_i[n]$ are as in part (a) to determine the coefficients a_i in terms of $x[n]$.

3.70. (a) In this problem, we consider the definition of the two-dimensional Fourier series for periodic signals with two independent variables. Specifically, consider a signal $x(t_1, t_2)$ that satisfies the equation

$$x(t_1, t_2) = x(t_1 + T_1, t_2 + T_2), \text{ for all } t_1, t_2.$$

This signal is periodic with period T_1 in the t_1 direction and with period T_2 in the t_2 direction. Such a signal has a series representation of the form

$$x(t_1, t_2) = \sum_{n=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} a_{mn} e^{j(m\omega_1 t_1 + n\omega_2 t_2)},$$

where

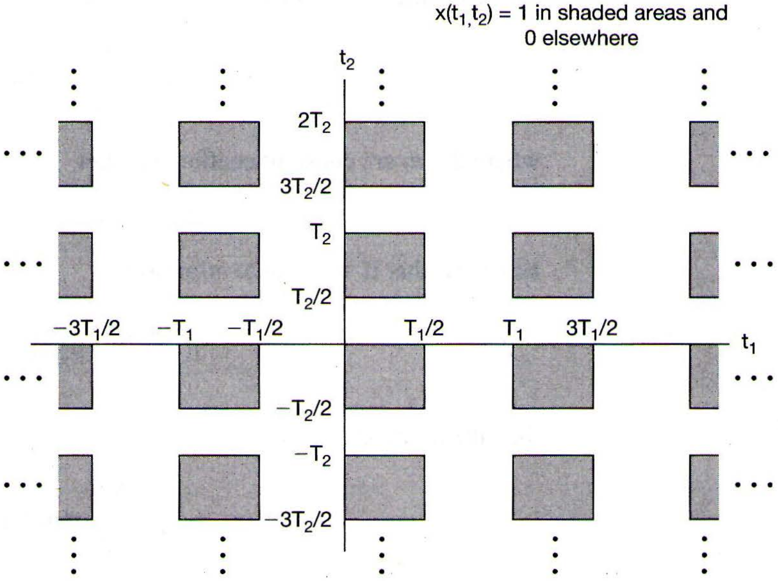
$$\omega_1 = 2\pi/T_1, \quad \omega_2 = 2\pi/T_2.$$

Find an expression for a_{mn} in terms of $x(t_1, t_2)$.

(b) Determine the Fourier series coefficients a_{mn} for the following signals:

(i) $\cos(2\pi t_1 + 2t_2)$

(ii) the signal illustrated in Figure P3.70



x(t1,t2) = 1 in shaded areas and 0 elsewhere
t2
2T2
3T2/2
T2
T2/2
−3T1/2
−T1
−T1/2
T1/2
T1
3T1/2
t1
−T2/2
−T2
−3T2/2

Figure P3.70

3.71. Consider the mechanical system shown in Figure P3.71. The differential equation relating velocity $v(t)$ and the input force $f(t)$ is given by

$$Bv(t) + K\int v(t)\,dt = f(t).$$

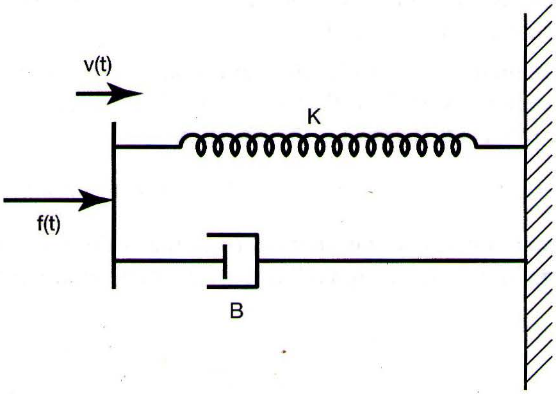


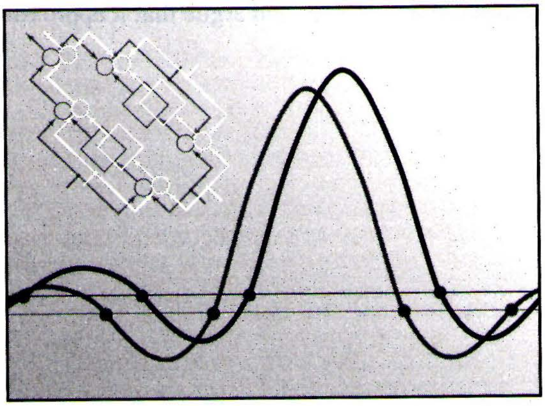
Figure P3.71

(a) Assuming that the output is $f_s(t)$, the compressive force acting on the spring, write the differential equation relating $f_s(t)$ and $f(t)$. Obtain the frequency response of the system, and argue that it approximates that of a lowpass filter.

(b) Assuming that the output is $f_d(t)$, the compressive force acting on the dashpot, write the differential equation relating $f_d(t)$ and $f(t)$. Obtain the frequency response of the system, and argue that it approximates that of a highpass filter.

4

THE CONTINUOUS-TIME FOURIER TRANSFORM



4.0 INTRODUCTION

In Chapter 3, we developed a representation of periodic signals as linear combinations of complex exponentials. We also saw how this representation can be used in describing the effect of LTI systems on signals.

In this and the following chapter, we extend these concepts to apply to signals that are not periodic. As we will see, a rather large class of signals, including all signals with finite energy, can also be represented through a linear combination of complex exponentials. Whereas for periodic signals the complex exponential building blocks are harmonically related, for aperiodic signals they are infinitesimally close in frequency, and the representation in terms of a linear combination takes the form of an integral rather than a sum. The resulting spectrum of coefficients in this representation is called the Fourier transform, and the synthesis integral itself, which uses these coefficients to represent the signal as a linear combination of complex exponentials, is called the inverse Fourier transform.

The development of this representation for aperiodic signals in continuous time is one of Fourier's most important contributions, and our development of the Fourier transform follows very closely the approach he used in his original work. In particular, Fourier reasoned that an aperiodic signal can be viewed as a periodic signal with an infinite period. More precisely, in the Fourier series representation of a periodic signal, as the period increases the fundamental frequency decreases and the harmonically related components become closer in frequency. As the period becomes infinite, the frequency components form a continuum and the Fourier series sum becomes an integral. In the next section we develop the Fourier series representation for continuous-time periodic signals, and in the sections that follow we build on this foundation as we explore many of the important properties of the continuous-time Fourier transform that form the foundation of frequency-domain methods for continuous-time signals and systems. In Chapter 5, we parallel this development for discrete-time signals.

4.1 REPRESENTATION OF APERIODIC SIGNALS: THE CONTINUOUS-TIME FOURIER TRANSFORM

4.1.1 Development of the Fourier Transform Representation of an Aperiodic Signal

To gain some insight into the nature of the Fourier transform representation, we begin by revisiting the Fourier series representation for the continuous-time periodic square wave examined in Example 3.5. Specifically, over one period,

$$x(t) = \begin{cases} 1, & |t| < T_1 \\ 0, & T_1 < |t| < T/2 \end{cases}$$

and periodically repeats with period T, as shown in Figure 4.1.

As determined in Example 3.5, the Fourier series coefficients a_k for this square wave are

[eq. (3.44)] $$a_k = \frac{2\sin(k\omega_0 T_1)}{k\omega_0 T}, \tag{4.1}$$

where $\omega_0 = 2\pi/T$. In Figure 3.7, bar graphs of these coefficients were shown for a fixed value of T_1 and several different values of T.

An alternative way of interpreting eq. (4.1) is as samples of an envelope function, specifically,

$$Ta_k = \left.\frac{2\sin\omega T_1}{\omega}\right|_{\omega = k\omega_0}. \tag{4.2}$$

That is, with ω thought of as a continuous variable, the function $(2\sin\omega T_1)/\omega$ represents the envelope of Ta_k, and the coefficients a_k are simply equally spaced samples of this envelope. Also, for fixed T_1, the envelope of Ta_k is independent of T. In Figure 4.2, we again show the Fourier series coefficients for the periodic square wave, but this time as samples of the envelope of Ta_k, as specified in eq. (4.2). From the figure, we see that as

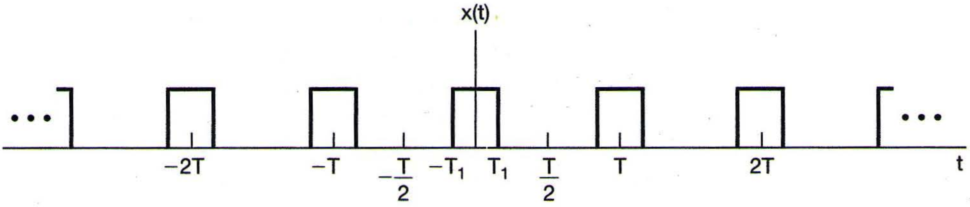


Figure 4.1 A continuous-time periodic square wave.

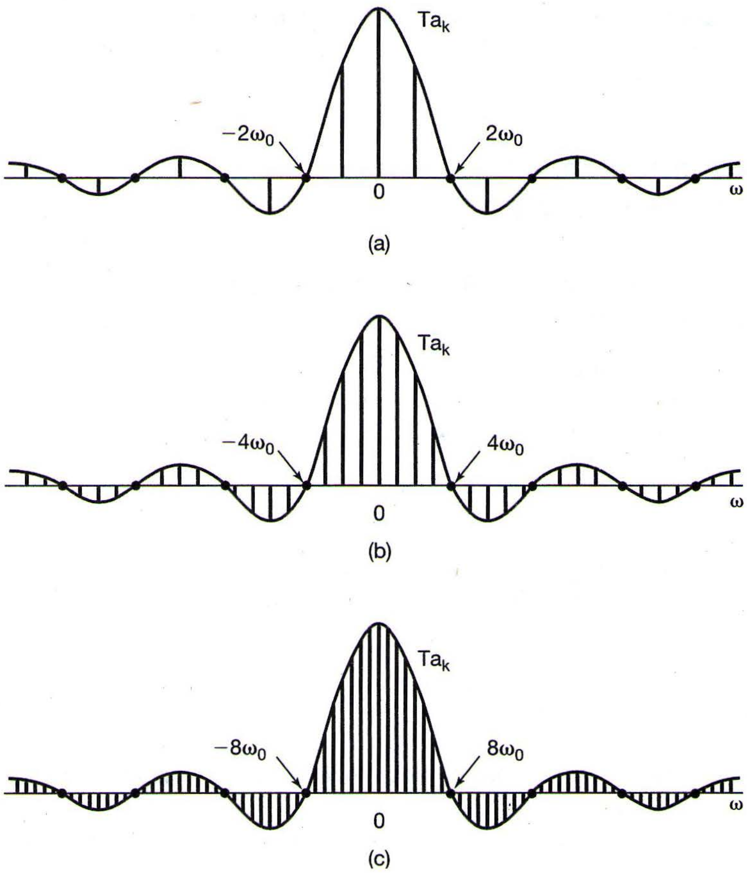


Figure 4.2 The Fourier series coefficients and their envelope for the periodic square wave in Figure 4.1 for several values of T (with T_1 fixed): (a) $T = 4T_1$; (b) $T = 8T_1$; (c) $T = 16T_1$.

T increases, or equivalently, as the fundamental frequency $\omega_0 = 2\pi/T$ decreases, the envelope is sampled with a closer and closer spacing. As T becomes arbitrarily large, the original periodic square wave approaches a rectangular pulse (i.e., all that remains in the time domain is an aperiodic signal corresponding to one period of the square wave). Also, the Fourier series coefficients, multiplied by T, become more and more closely spaced samples of the envelope, so that in some sense (which we will specify shortly) the set of Fourier series coefficients approaches the envelope function as $T \rightarrow \infty$.

This example illustrates the basic idea behind Fourier's development of a representation for aperiodic signals. Specifically, we think of an aperiodic signal as the limit of a periodic signal as the period becomes arbitrarily large, and we examine the limiting behavior of the Fourier series representation for this signal. In particular, consider a signal $x(t)$ that is of finite duration. That is, for some number T_1, $x(t) = 0$ if $|t| > T_1$, as illustrated in Figure 4.3(a). From this aperiodic signal, we can construct a periodic signal $\tilde{x}(t)$ for which $x(t)$ is one period, as indicated in Figure 4.3(b). As we choose the period T to be larger, $\tilde{x}(t)$ is identical to $x(t)$ over a longer interval, and as $T \rightarrow \infty$, $\tilde{x}(t)$ is equal to $x(t)$ for any finite value of t.

Let us now examine the effect of this on the Fourier series representation of $\tilde{x}(t)$. Rewriting eqs. (3.38) and (3.39) here for convenience, with the integral in eq. (3.39)

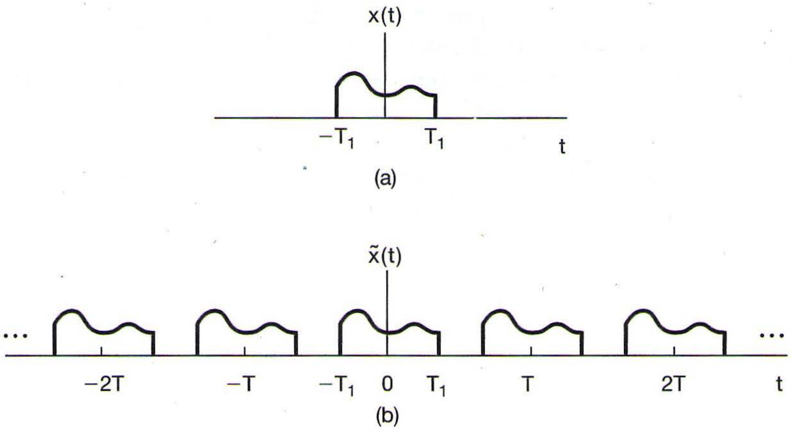


Figure 4.3 (a) Aperiodic signal $x(t)$; (b) periodic signal $\tilde{x}(t)$, constructed to be equal to $x(t)$ over one period.

carried out over the interval $-T/2 \le t \le T/2$, we have

$$\tilde{x}(t) = \sum_{k=-\infty}^{+\infty} a_k e^{jk\omega_0 t}, \tag{4.3}$$

$$a_k = \frac{1}{T}\int_{-T/2}^{T/2} \tilde{x}(t)e^{-jk\omega_0 t}dt, \tag{4.4}$$

where $\omega_0 = 2\pi/T$. Since $\tilde{x}(t) = x(t)$ for $|t| < T/2$, and also, since $x(t) = 0$ outside this interval, eq. (4.4) can be rewritten as

$$a_k = \frac{1}{T}\int_{-T/2}^{T/2} x(t)e^{-jk\omega_0 t}dt = \frac{1}{T}\int_{-\infty}^{+\infty} x(t)e^{-jk\omega_0 t}dt.$$

Therefore, defining the envelope $X(j\omega)$ of Ta_k as

$$X(j\omega) = \int_{-\infty}^{+\infty} x(t)e^{-j\omega t}dt, \tag{4.5}$$

we have, for the coefficients a_k,

$$a_k = \frac{1}{T}X(jk\omega_0). \tag{4.6}$$

Combining eqs. (4.6) and (4.3), we can express $\tilde{x}(t)$ in terms of $X(j\omega)$ as

$$\tilde{x}(t) = \sum_{k=-\infty}^{+\infty} \frac{1}{T}X(jk\omega_0)e^{jk\omega_0 t},$$

or equivalently, since $2\pi/T = \omega_0$,

$$\tilde{x}(t) = \frac{1}{2\pi}\sum_{k=-\infty}^{+\infty} X(jk\omega_0)e^{jk\omega_0 t}\omega_0. \tag{4.7}$$

As $T \to \infty$, $\tilde{x}(t)$ approaches $x(t)$, and consequently, in the limit eq. (4.7) becomes a representation of $x(t)$. Furthermore, $\omega_0 \to 0$ as $T \to \infty$, and the right-hand side of eq. (4.7) passes to an integral. This can be seen by considering the graphical interpretation of the equation, illustrated in Figure 4.4. Each term in the summation on the right-hand side is the area of a rectangle of height $X(jk\omega_0)e^{jk\omega_0 t}$ and width ω_0. (Here, t is regarded as fixed.) As $\omega_0 \to 0$, the summation converges to the integral of $X(j\omega)e^{j\omega t}$. Therefore, using the fact that $\tilde{x}(t) \to x(t)$ as $T \to \infty$, we see that eqs. (4.7) and (4.5) respectively become

$$\boxed{x(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} X(j\omega)e^{j\omega t}d\omega} \tag{4.8}$$

and

$$\boxed{X(j\omega) = \int_{-\infty}^{+\infty} x(t)e^{-j\omega t}dt.} \tag{4.9}$$

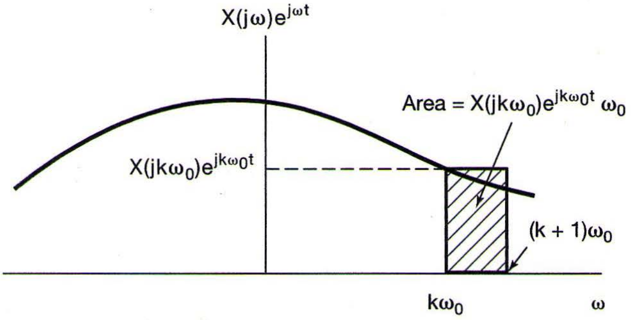

$X(j\omega)e^{j\omega t}$
Area = $X(jk\omega_0)e^{jk\omega_0 t}\ \omega_0$
$X(jk\omega_0)e^{jk\omega_0 t}$
$(k + 1)\omega_0$
$k\omega_0$
ω

Figure 4.4 Graphical interpretation of eq. (4.7).

Equations (4.8) and (4.9) are referred to as the *Fourier transform pair,* with the function $X(j\omega)$ referred to as the *Fourier Transform* or *Fourier integral* of $x(t)$ and eq. (4.8) as the *inverse Fourier transform* equation. The *synthesis* equation (4.8) plays a role for aperiodic signals similar to that of eq. (3.38) for periodic signals, since both represent a signal as a linear combination of complex exponentials. For periodic signals, these complex exponentials have amplitudes $\{a_k\}$, as given by eq. (3.39), and occur at a discrete set of harmonically related frequencies $k\omega_0$, $k = 0, \pm 1, \pm 2, \ldots$. For aperiodic signals, the complex exponentials occur at a continuum of frequencies and, according to the synthesis equation (4.8), have "amplitude" $X(j\omega)(d\omega/2\pi)$. In analogy with the terminology used for the Fourier series coefficients of a periodic signal, the transform $X(j\omega)$ of an aperiodic signal $x(t)$ is commonly referred to as the *spectrum* of $x(t)$, as it provides us with the information needed for describing $x(t)$ as a linear combination (specifically, an integral) of sinusoidal signals at different frequencies.

Based on the above development, or equivalently on a comparison of eq. (4.9) and eq. (3.39), we also note that the Fourier coefficients a_k of a periodic signal $\tilde{x}(t)$ can be expressed in terms of equally spaced *samples* of the Fourier transform of one period of $\tilde{x}(t)$. Specifically, suppose that $\tilde{x}(t)$ is a periodic signal with period T and Fourier coefficients

a_k. Let $x(t)$ be a finite-duration signal that is equal to $\tilde{x}(t)$ over exactly one period—say, for $s \leq t \leq s+T$ for some value of s—and that is zero otherwise. Then, since eq. (3.39) allows us to compute the Fourier coefficients of $\tilde{x}(t)$ by integrating over any period, we can write

$$a_k = \frac{1}{T}\int_s^{s+T} \tilde{x}(t)e^{-jk\omega_0 t}dt = \frac{1}{T}\int_s^{s+T} x(t)e^{-jk\omega_0 t}dt.$$

Since $x(t)$ is zero outside the range $s \leq t \leq s+T$ we can equivalently write

$$a_k = \frac{1}{T}\int_{-\infty}^{+\infty} x(t)e^{-jk\omega_0 t}dt.$$

Comparing with eq. (4.9) we conclude that

$$a_k = \frac{1}{T}X(j\omega)\bigg|_{\omega = k\omega_0}, \tag{4.10}$$

where $X(j\omega)$ is the Fourier transform of $x(t)$. Equation 4.10 states that the Fourier coefficients of $\tilde{x}(t)$ are proportional to samples of the Fourier transform of one period of $\tilde{x}(t)$. This fact, which is often of use in practice, is examined further in Problem 4.37.

4.1.2 Convergence of Fourier Transforms

Although the argument we used in deriving the Fourier transform pair assumed that $x(t)$ was of arbitrary but finite duration, eqs. (4.8) and (4.9) remain valid for an extremely broad class of signals of infinite duration. In fact, our derivation of the Fourier transform suggests that a set of conditions like those required for the convergence of Fourier series should also apply here, and indeed, that can be shown to be the case.[1] Specifically, consider $X(j\omega)$ evaluated according to eq. (4.9), and let $\hat{x}(t)$ denote the signal obtained by using $X(j\omega)$ in the right-hand side of eq. (4.8). That is,

$$\hat{x}(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} X(j\omega)e^{j\omega t}d\omega.$$

What we would like to know is when eq. (4.8) is valid [i.e., when is $\hat{x}(t)$ a valid representation of the original signal $x(t)$?]. If $x(t)$ has finite energy, i.e., if it is square integrable, so that

$$\int_{-\infty}^{+\infty} |x(t)|^2 dt < \infty, \tag{4.11}$$

then we are guaranteed that $X(j\omega)$ is finite [i.e., eq. (4.9) converges] and that, with $e(t)$ denoting the error between $\hat{x}(t)$ and $x(t)$ [i.e., $e(t) = \hat{x}(t) - x(t)$],

---

[1] For a mathematically rigorous discussion of the Fourier transform and its properties and applications, see R. Bracewell, *The Fourier Transform and Its Applications,* 2nd ed. (New York: McGraw-Hill Book Company, 1986); A. Papoulis, *The Fourier Integral and Its Applications* (New York: McGraw-Hill Book Company, 1987); E. C. Titchmarsh, *Introduction to the Theory of Fourier Integrals* (Oxford: Clarendon Press, 1948); and the book by Dym and McKean referenced in footnote 2 of Chapter 3.

$$\int_{-\infty}^{+\infty} |e(t)|^2 dt = 0. \tag{4.12}$$

Equations (4.11) and (4.12) are the aperiodic counterparts of eqs. (3.51) and (3.54) for periodic signals. Thus, in a manner similar to that for periodic signals, if $x(t)$ has finite energy, then, although $x(t)$ and its Fourier representation $\hat{x}(t)$ may differ significantly at individual values of t, there is no energy in their difference.

Just as with periodic signals, there is an alternative set of conditions which are sufficient to ensure that $\hat{x}(t)$ is equal to $x(t)$ for any t except at a discontinuity, where it is equal to the average of the values on either side of the discontinuity. These conditions, again referred to as the Dirichlet conditions, require that:

1. $x(t)$ be absolutely integrable; that is,

$$\int_{-\infty}^{+\infty} |x(t)| dt < \infty. \tag{4.13}$$

2. $x(t)$ have a finite number of maxima and minima within any finite interval.
3. $x(t)$ have a finite number of discontinuities within any finite interval. Futhermore, each of these discontinuities must be finite.

Therefore, absolutely integrable signals that are continuous or that have a finite number of discontinuities have Fourier transforms.

Although the two alternative sets of conditions that we have given are sufficient to guarantee that a signal has a Fourier transform, we will see in the next section that periodic signals, which are neither absolutely integrable nor square integrable over an *infinite* interval, can be considered to have Fourier transforms if impulse functions are permitted in the transform. This has the advantage that the Fourier series and Fourier transform can be incorporated into a common framework, which we will find to be very convenient in subsequent chapters. Before examining the point further in Section 4.2, however, let us consider several examples of the Fourier transform.

4.1.3 Examples of Continuous-Time Fourier Transforms

Example 4.1

Consider the signal

$$x(t) = e^{-at}u(t) \quad a > 0.$$

From eq. (4.9),

$$X(j\omega) = \int_0^{\infty} e^{-at} e^{-j\omega t} dt = -\frac{1}{a + j\omega} e^{-(a+j\omega)t} \Big|_0^{\infty}.$$

That is,

$$X(j\omega) = \frac{1}{a + j\omega}, \quad a > 0.$$

Since this Fourier transform is complex valued, to plot it as a function of ω, we express $X(j\omega)$ in terms of its magnitude and phase:

$$|X(j\omega)| = \frac{1}{\sqrt{a^2+\omega^2}}, \qquad \sphericalangle X(j\omega) = -\tan^{-1}\left(\frac{\omega}{a}\right).$$

Each of these components is sketched in Figure 4.5.

Note that if a is complex rather than real, then $x(t)$ is absolutely integrable as long as $\Re e\{a\} > 0$, and in this case the preceding calculation yields the same form for $X(j\omega)$. That is,

$$X(j\omega) = \frac{1}{a+j\omega}, \qquad \Re e\{a\} > 0.$$

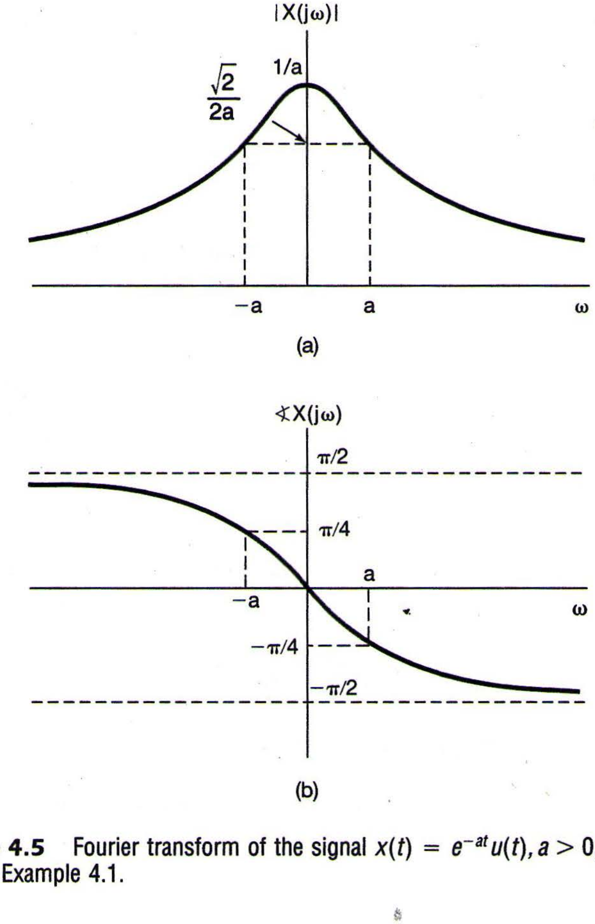


Figure 4.5 Fourier transform of the signal $x(t) = e^{-at}u(t), a > 0$, considered in Example 4.1.

Example 4.2

Let

$$x(t) = e^{-a|t|}, \quad a > 0.$$

This signal is sketched in Figure 4.6. The Fourier transform of the signal is

$$\begin{aligned} X(j\omega) &= \int_{-\infty}^{+\infty} e^{-a|t|}e^{-j\omega t}\,dt = \int_{-\infty}^{0} e^{at}e^{-j\omega t}\,dt + \int_{0}^{\infty} e^{-at}e^{-j\omega t}\,dt \\ &= \frac{1}{a-j\omega} + \frac{1}{a+j\omega} \\ &= \frac{2a}{a^2+\omega^2}. \end{aligned}$$

In this case $X(j\omega)$ is real, and it is illustrated in Figure 4.7.

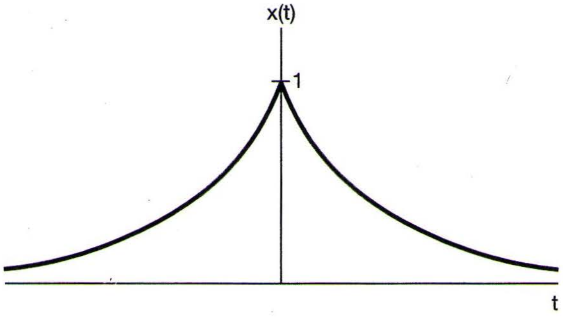


Figure 4.6 Signal $x(t) = e^{-a|t|}$ of Example 4.2.

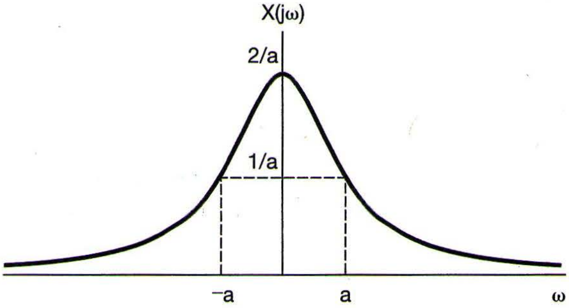


Figure 4.7 Fourier transform of the signal considered in Example 4.2 and depicted in Figure 4.6.

Example 4.3

Now let us determine the Fourier transform of the unit impulse

$$x(t) = \delta(t). \tag{4.14}$$

Substituting into eq. (4.9) yields

$$X(j\omega) = \int_{-\infty}^{+\infty} \delta(t)e^{-j\omega t}dt = 1. \tag{4.15}$$

That is, the unit impulse has a Fourier transform consisting of equal contributions at *all* frequencies.

Example 4.4

Consider the rectangular pulse signal

$$x(t) = \begin{cases} 1, & |t| < T_1 \\ 0, & |t| > T_1 \end{cases}, \tag{4.16}$$

as shown in Figure 4.8(a). Applying eq. (4.9), we find that the Fourier transform of this signal is

$$X(j\omega) = \int_{-T_1}^{T_1} e^{-j\omega t}\,dt = 2\frac{\sin \omega T_1}{\omega}, \tag{4.17}$$

as sketched in Figure 4.8(b).

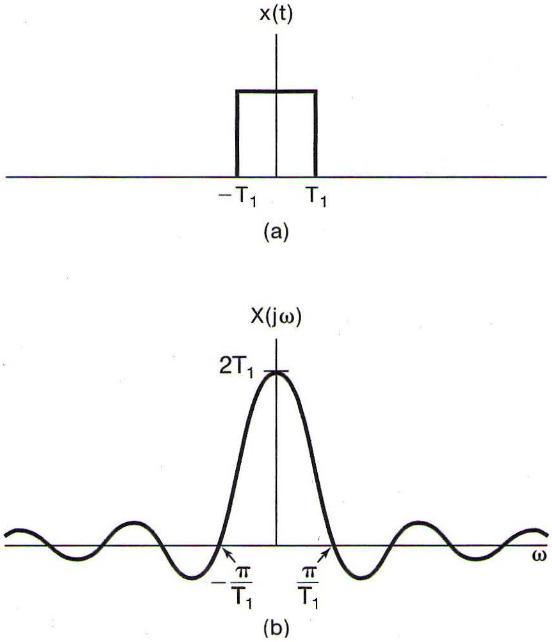


Figure 4.8 (a) The rectangular pulse signal of Example 4.4 and (b) its Fourier transform.

As we discussed at the beginning of this section, the signal given by eq. (4.16) can be thought of as the limiting form of a periodic square wave as the period becomes arbitrarily large. Therefore, we might expect that the convergence of the synthesis equation for this signal would behave in a manner similar to that observed in Example 3.5 for the square wave. This is, in fact, the case. Specifically, consider the inverse Fourier transform for the rectangular pulse signal:

$$\hat{x}(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} 2\frac{\sin \omega T_1}{\omega} e^{j\omega t} d\omega.$$

Then, since $x(t)$ is square integrable,

$$\int_{-\infty}^{+\infty} |x(t) - \hat{x}(t)|^2\, dt = 0.$$

Furthermore, because $x(t)$ satisfies the Dirichlet conditions, $\hat{x}(t) = x(t)$, except at the points of discontinuity, $t = \pm T_1$, where $\hat{x}(t)$ converges to 1/2, which is the average of the values of $x(t)$ on both sides of the discontinuity. In addition, the convergence of $\hat{x}(t)$ to $x(t)$ exhibits the Gibbs phenomenon, much as was illustrated for the periodic square wave in Figure 3.9. Specifically, in analogy with the finite Fourier series approximation, eq. (3.47), consider the following integral over a finite-length interval of frequencies:

$$\frac{1}{2\pi}\int_{-W}^{W} 2\frac{\sin \omega T_1}{\omega} e^{j\omega t}\, d\omega.$$

As $W \to \infty$, this signal converges to $x(t)$ everywhere, except at the discontinuities. Moreover, the signal exhibits ripples near the discontinuities. The peak amplitude of these ripples does not decrease as W increases, although the ripples do become compressed toward the discontinuity, and the energy in the ripples converges to zero.

Example 4.5

Consider the signal $x(t)$ whose Fourier transform is

$$X(j\omega) = \begin{cases} 1, & |\omega| < W \\ 0, & |\omega| > W \end{cases}. \tag{4.18}$$

This transform is illustrated in Figure 4.9(a). Using the synthesis equation (4.8), we can

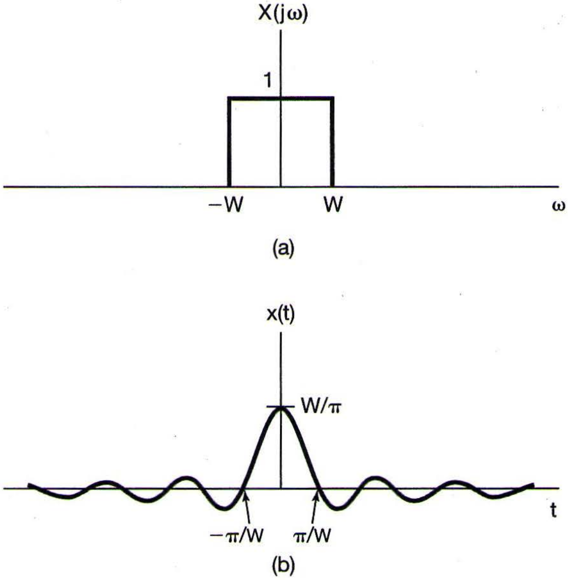


Figure 4.9 Fourier transform pair of Example 4.5: (a) Fourier transform for Example 4.5 and (b) the corresponding time function.

then determine

$$x(t) = \frac{1}{2\pi}\int_{-W}^{W} e^{j\omega t}d\omega = \frac{\sin Wt}{\pi t}, \tag{4.19}$$

which is depicted in Figure 4.9(b).

Comparing Figures 4.8 and 4.9 or, equivalently, eqs. (4.16) and (4.17) with eqs. (4.18) and (4.19), we see an interesting relationship. In each case, the Fourier transform pair consists of a function of the form $(\sin a\theta)/b\theta$ and a rectangular pulse. However, in Example 4.4, it is the *signal* $x(t)$ that is a pulse, while in Example 4.5, it is the *transform* $X(j\omega)$. The special relationship that is apparent here is a direct consequence of the *duality property* for Fourier transforms, which we discuss in detail in Section 4.3.6.

Functions of the form given in eqs. (4.17) and (4.19) arise frequently in Fourier analysis and in the study of LTI systems and are referred to as *sinc functions.* A commonly used precise form for the sinc function is

$$\operatorname{sinc}(\theta) = \frac{\sin \pi\theta}{\pi\theta}. \tag{4.20}$$

The sinc function is plotted in Figure 4.10. Both of the signals in eqs. (4.17) and (4.19) can be expressed in terms of the sinc function:

$$\frac{2\sin\omega T_1}{\omega} = 2T_1 \operatorname{sinc}\left(\frac{\omega T_1}{\pi}\right)$$

$$\frac{\sin Wt}{\pi t} = \frac{W}{\pi}\operatorname{sinc}\left(\frac{Wt}{\pi}\right).$$

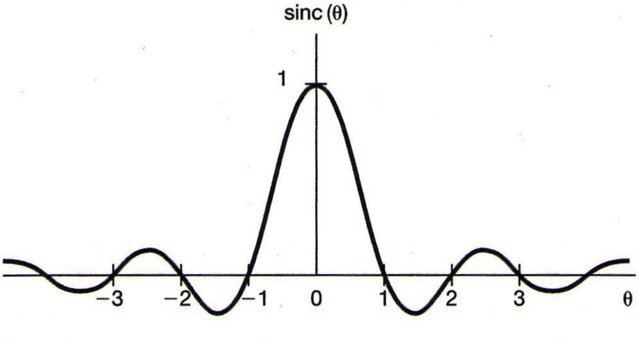


Figure 4.10 The sinc function.

Finally, we can gain insight into one other property of the Fourier transform by examining Figure 4.9, which we have redrawn as Figure 4.11 for several different values of W. From this figure, we see that as W increases, $X(j\omega)$ becomes broader, while the main peak of $x(t)$ at $t = 0$ becomes higher and the width of the first lobe of this signal (i.e., the part of the signal for $|t| < \pi/W$) becomes narrower. In fact, in the limit as $W \to \infty$, $X(j\omega) = 1$ for all ω, and consequently, from Example 4.3, we see that $x(t)$ in eq. (4.19) converges to an impulse as $W \to \infty$. The behavior depicted in Figure 4.11 is an example of the inverse relationship that exists between the time and frequency domains,

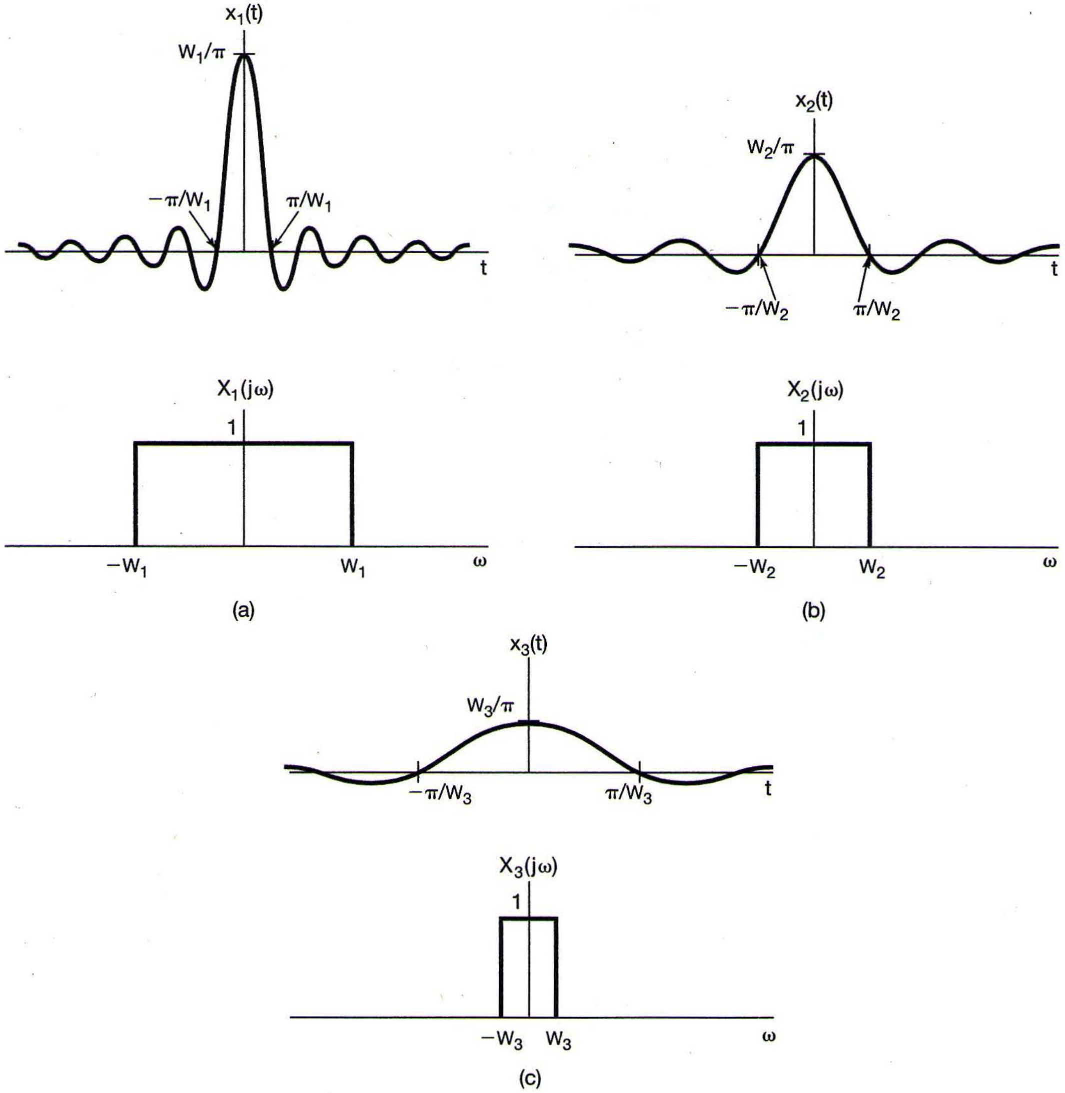


Figure 4.11 Fourier transform pair of Figure 4.9 for several different values of *W*.

and we can see a similar effect in Figure 4.8, where an increase in T_1 broadens $x(t)$ but makes $X(j\omega)$ narrower. In Section 4.3.5, we provide an explanation of this behavior in the context of the scaling property of the Fourier transform.

4.2 THE FOURIER TRANSFORM FOR PERIODIC SIGNALS

In the preceding section, we introduced the Fourier transform representation and gave several examples. While our attention in that section was focused on aperiodic signals, we can also develop Fourier transform representations for periodic signals, thus allowing us to consider both periodic and aperiodic signals within a unified context. In fact, as we will see, we can construct the Fourier transform of a periodic signal directly from its Fourier series representation. The resulting transform consists of a train of impulses in the frequency domain, with the areas of the impulses proportional to the Fourier series coefficients. This will turn out to be a very useful representation.

To suggest the general result, let us consider a signal $x(t)$ with Fourier transform $X(j\omega)$ that is a single impulse of area 2π at $\omega = \omega_0$; that is,

$$X(j\omega) = 2\pi\delta(\omega - \omega_0). \tag{4.21}$$

To determine the signal $x(t)$ for which this is the Fourier transform, we can apply the inverse transform relation, eq. (4.8), to obtain

$$\begin{aligned} x(t) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} 2\pi\delta(\omega - \omega_0)e^{j\omega t}\,d\omega \\ &= e^{j\omega_0 t}. \end{aligned}$$

More generally, if $X(j\omega)$ is of the form of a linear combination of impulses equally spaced in frequency, that is,

$$X(j\omega) = \sum_{k=-\infty}^{+\infty} 2\pi a_k \delta(\omega - k\omega_0), \tag{4.22}$$

then the application of eq. (4.8) yields

$$x(t) = \sum_{k=-\infty}^{+\infty} a_k e^{jk\omega_0 t}. \tag{4.23}$$

We see that eq. (4.23) corresponds exactly to the Fourier *series* representation of a periodic signal, as specified by eq. (3.38). Thus, the Fourier transform of a periodic signal with Fourier series coefficients $\{a_k\}$ can be interpreted as a train of impulses occurring at the harmonically related frequencies and for which the area of the impulse at the kth harmonic frequency $k\omega_0$ is 2π times the kth Fourier series coefficient a_k.

Example 4.6

Consider again the square wave illustrated in Figure 4.1. The Fourier series coefficients for this signal are

$$a_k = \frac{\sin k\omega_0 T_1}{\pi k},$$

and the Fourier transform of the signal is

$$X(j\omega) = \sum_{k=-\infty}^{+\infty} \frac{2\sin k\omega_0 T_1}{k}\delta(\omega - k\omega_0),$$

which is sketched in Figure 4.12 for $T = 4T_1$. In comparison with Figure 3.7(a), the only differences are a proportionality factor of 2π and the use of impulses rather than a bar graph.

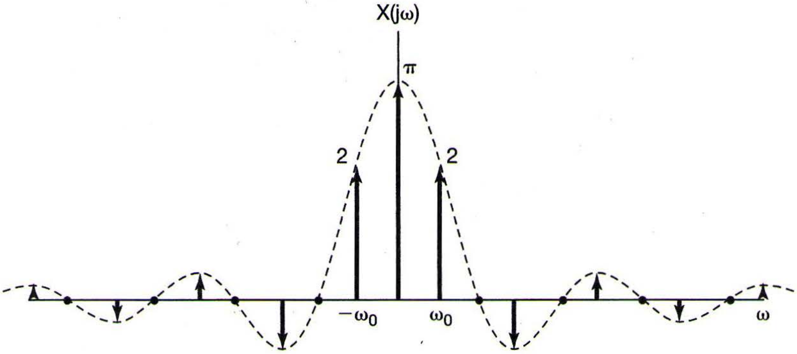


Figure 4.12 Fourier transform of a symmetric periodic square wave.

Example 4.7

Let

$$x(t) = \sin \omega_0 t.$$

The Fourier series coefficients for this signal are

$$a_1 = \frac{1}{2j},$$

$$a_{-1} = -\frac{1}{2j},$$

$$a_k = 0, \qquad k \neq 1 \quad \text{or} \quad -1.$$

Thus, the Fourier transform is as shown in Figure 4.13(a). Similarly, for

$$x(t) = \cos \omega_0 t,$$

the Fourier series coefficients are

$$a_1 = a_{-1} = \frac{1}{2},$$

$$a_k = 0, \qquad k \neq 1 \quad \text{or} \quad -1.$$

The Fourier transform of this signal is depicted in Figure 4.13(b). These two transforms will be of considerable importance when we analyze sinusoidal modulation systems in Chapter 8.

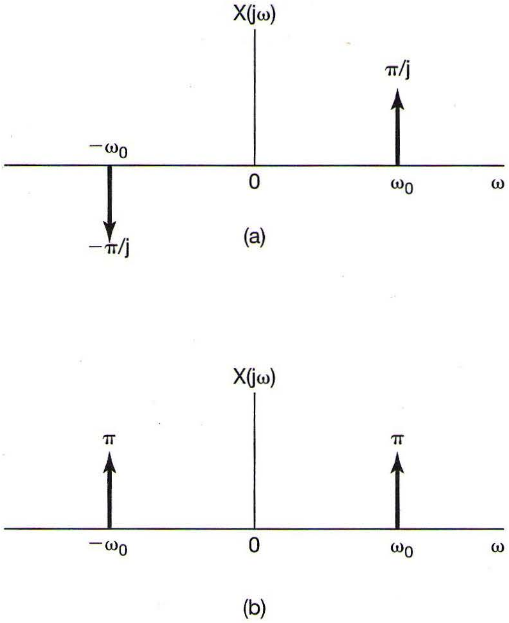


Figure 4.13 Fourier transforms of (a) $x(t) = \sin \omega_0 t$; (b) $x(t) = \cos \omega_0 t$.

Example 4.8

A signal that we will find extremely useful in our analysis of sampling systems in Chapter 7 is the impulse train

$$x(t) = \sum_{k=-\infty}^{+\infty} \delta(t - kT),$$

which is periodic with period T, as indicated in Figure 4.14(a). The Fourier series coefficients for this signal were computed in Example 3.8 and are given by

$$a_k = \frac{1}{T}\int_{-T/2}^{+T/2} \delta(t) e^{-jk\omega_0 t}\, dt = \frac{1}{T}.$$

That is, every Fourier coefficient of the periodic impulse train has the same value, $1/T$. Substituting this value for a_k in eq. (4.22) yields

$$X(j\omega) = \frac{2\pi}{T} \sum_{k=-\infty}^{+\infty} \delta\left(\omega - \frac{2\pi k}{T}\right).$$

Thus, the Fourier transform of a periodic impulse train in the time domain with period T is a periodic impulse train in the frequency domain with period $2\pi/T$, as sketched in Figure 4.14(b). Here again, we see an illustration of the inverse relationship between the time and the frequency domains. As the spacing between the impulses in the time domain (i.e., the period) gets longer, the spacing between the impulses in the frequency domain (namely, the fundamental frequency) gets smaller.

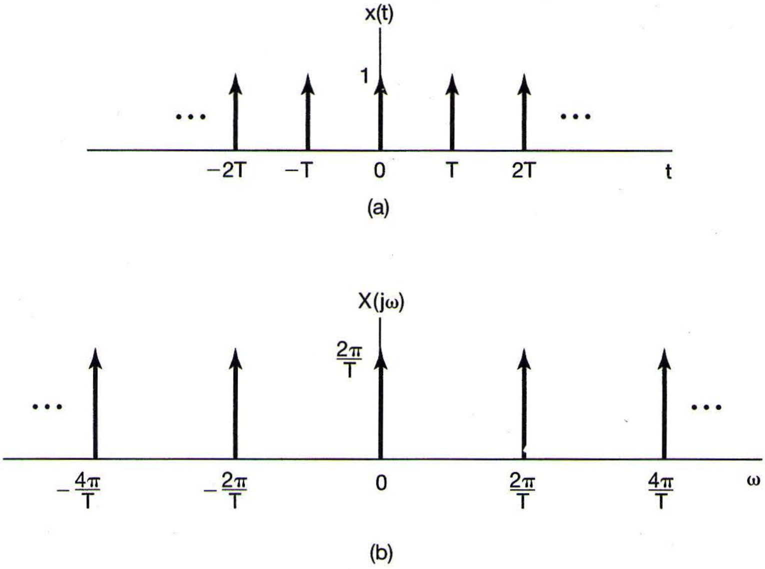


Figure 4.14 (a) Periodic impulse train; (b) its Fourier transform.

4.3 PROPERTIES OF THE CONTINUOUS-TIME FOURIER TRANSFORM

In this and the following two sections, we consider a number of properties of the Fourier transform. A detailed listing of these properties is given in Table 4.1 in Section 4.6. As was the case for the Fourier series representation of periodic signals, these properties provide us with a significant amount of insight into the transform and into the relationship between the time-domain and frequency-domain descriptions of a signal. In addition, many of the properties are often useful in reducing the complexity of the evaluation of Fourier transforms or inverse transforms. Furthermore, as described in the preceding section, there is a close relationship between the Fourier series and Fourier transform representations of a periodic signal, and using this relationship, we can translate many of the Fourier transform properties into corresponding Fourier series properties, which we discussed independently in Chapter 3. (See, in particular, Section 3.5 and Table 3.1.)

Throughout the discussion in this section, we will be referring frequently to functions of time and their Fourier transforms, and we will find it convenient to use a shorthand notation to indicate the pairing of a signal and its transform. As developed in Section 4.1, a signal $x(t)$ and its Fourier transform $X(j\omega)$ are related by the Fourier transform synthesis and analysis equations,

[eq. (4.8)] $$x(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} X(j\omega)e^{j\omega t}d\omega \tag{4.24}$$

and

[eq. (4.9)] $$X(j\omega) = \int_{-\infty}^{+\infty} x(t)e^{-j\omega t}dt. \tag{4.25}$$

We will sometimes find it convenient to refer to $X(j\omega)$ with the notation $\mathfrak{F}\{x(t)\}$ and to $x(t)$ with the notation $\mathfrak{F}^{-1}\{X(j\omega)\}$. We will also refer to $x(t)$ and $X(j\omega)$ as a Fourier transform pair with the notation

$$x(t) \overset{\mathfrak{F}}{\longleftrightarrow} X(j\omega).$$

Thus, with reference to Example 4.1,

$$\frac{1}{a + j\omega} = \mathfrak{F}\{e^{-at}u(t)\},$$

$$e^{-at}u(t) = \mathfrak{F}^{-1}\left\{\frac{1}{a + j\omega}\right\},$$

and

$$e^{-at}u(t) \overset{\mathfrak{F}}{\longleftrightarrow} \frac{1}{a + j\omega}.$$

4.3.1 Linearity

If

$$x(t) \overset{\mathfrak{F}}{\longleftrightarrow} X(j\omega)$$

and

$$y(t) \overset{\mathfrak{F}}{\longleftrightarrow} Y(j\omega),$$

then

$$\boxed{ax(t) + by(t) \overset{\mathfrak{F}}{\longleftrightarrow} aX(j\omega) + bY(j\omega).} \qquad (4.26)$$

The proof of eq. (4.26) follows directly by application of the analysis eq. (4.25) to $ax(t) + by(t)$. The linearity property is easily extended to a linear combination of an arbitrary number of signals.

4.3.2 Time Shifting

If

$$x(t) \overset{\mathfrak{F}}{\longleftrightarrow} X(j\omega),$$

then

$$\boxed{x(t - t_0) \overset{\mathfrak{F}}{\longleftrightarrow} e^{-j\omega t_0}X(j\omega).} \qquad (4.27)$$

To establish this property, consider eq. (4.24):

$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(j\omega)e^{j\omega t}d\omega.$$

Replacing t by $t - t_0$ in this equation, we obtain

$$\begin{aligned} x(t-t_0) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} X(j\omega)e^{j\omega(t-t_0)}d\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left(e^{-j\omega t_0}X(j\omega)\right)e^{j\omega t}d\omega. \end{aligned}$$

Recognizing this as the synthesis equation for $x(t - t_0)$, we conclude that

$$\mathfrak{F}\{x(t-t_0)\} = e^{-j\omega t_0}X(j\omega).$$

One consequence of the time-shift property is that a signal which is shifted in time does not have the *magnitude* of its Fourier transform altered. That is, if we express $X(j\omega)$ in polar form as

$$\mathfrak{F}\{x(t)\} = X(j\omega) = |X(j\omega)|e^{j\sphericalangle X(j\omega)},$$

then

$$\mathfrak{F}\{x(t-t_0)\} = e^{-j\omega t_0}X(j\omega) = |X(j\omega)|e^{j[\sphericalangle X(j\omega)-\omega t_0]}.$$

Thus, the effect of a time shift on a signal is to introduce into its transform a phase shift, namely, $-\omega t_0$, which is a linear function of ω.

Example 4.9

To illustrate the usefulness of the Fourier transform linearity and time-shift properties, let us consider the evaluation of the Fourier transform of the signal $x(t)$ shown in Figure 4.15(a).

First, we observe that $x(t)$ can be expressed as the linear combination

$$x(t) = \frac{1}{2}x_1(t-2.5) + x_2(t-2.5),$$

where the signals $x_1(t)$ and $x_2(t)$ are the rectangular pulse signals shown in Figure 4.15(b) and (c). Then, using the result from Example 4.4, we obtain

$$X_1(j\omega) = \frac{2\sin(\omega/2)}{\omega} \quad \text{and} \quad X_2(j\omega) = \frac{2\sin(3\omega/2)}{\omega}.$$

Finally, using the linearity and time-shift properties of the Fourier transform yields

$$X(j\omega) = e^{-j5\omega/2}\left\{\frac{\sin(\omega/2) + 2\sin(3\omega/2)}{\omega}\right\}.$$

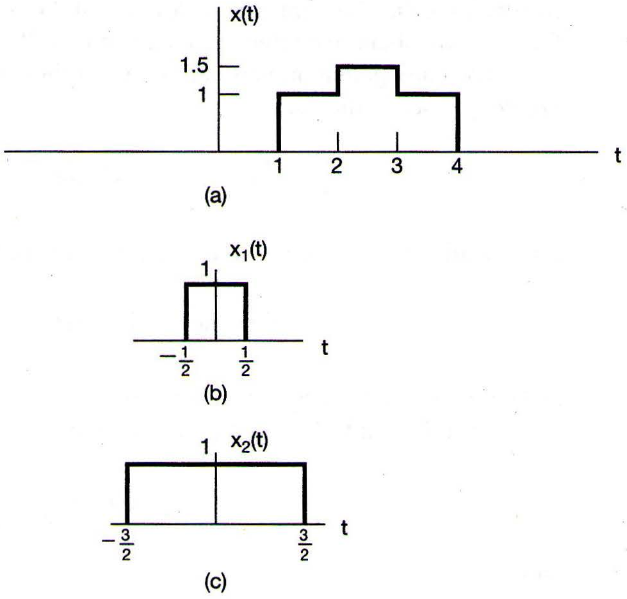


Figure 4.15 Decomposing a signal into the linear combination of two simpler signals. (a) The signal $x(t)$ for Example 4.9; (b) and (c) the two component signals used to represent $x(t)$.

4.3.3 Conjugation and Conjugate Symmetry

The conjugation property states that if

$$x(t) \overset{\mathfrak{F}}{\longleftrightarrow} X(j\omega),$$

then

$$\boxed{x^*(t) \overset{\mathfrak{F}}{\longleftrightarrow} X^*(-j\omega).} \tag{4.28}$$

This property follows from the evaluation of the complex conjugate of eq. (4.25):

$$\begin{aligned} X^*(j\omega) &= \left[\int_{-\infty}^{+\infty} x(t)e^{-j\omega t}\,dt\right]^* \\ &= \int_{-\infty}^{+\infty} x^*(t)e^{j\omega t}\,dt. \end{aligned}$$

Replacing ω by $-\omega$, we see that

$$X^*(-j\omega) = \int_{-\infty}^{+\infty} x^*(t)e^{-j\omega t}\,dt. \tag{4.29}$$

Recognizing that the right-hand side of eq. (4.29) is the Fourier transform analysis equation for $x^*(t)$, we obtain the relation given in eq. (4.28).

The conjugation property allows us to show that if $x(t)$ is real, then $X(j\omega)$ has *conjugate symmetry;* that is,

$$\boxed{X(-j\omega) = X^*(j\omega) \qquad [x(t) \text{ real}].} \tag{4.30}$$

Specifically, if $x(t)$ is real so that $x^*(t) = x(t)$, we have, from eq. (4.29),

$$X^*(-j\omega) = \int_{-\infty}^{+\infty} x(t)e^{j\omega t}\,dt = X(j\omega),$$

and eq. (4.30) follows by replacing ω with $-\omega$.

From Example 4.1, with $x(t) = e^{-at}u(t)$,

$$X(j\omega) = \frac{1}{a + j\omega}$$

and

$$X(-j\omega) = \frac{1}{a - j\omega} = X^*(j\omega).$$

As one consequence of eq. (4.30), if we express $X(j\omega)$ in rectangular form as

$$X(j\omega) = \mathcal{R}e\{X(j\omega)\} + j\mathcal{I}m\{X(j\omega)\},$$

then if $x(t)$ is real,

$$\mathcal{R}e\{X(j\omega)\} = \mathcal{R}e\{X(-j\omega)\}$$

and

$$\mathcal{I}m\{X(j\omega)\} = -\mathcal{I}m\{X(-j\omega)\}.$$

That is, the real part of the Fourier transform is an *even* function of frequency, and the imaginary part is an *odd* function of frequency. Similarly, if we express $X(j\omega)$ in polar form as

$$X(j\omega) = |X(j\omega)|e^{j\sphericalangle X(j\omega)},$$

then it follows from eq. (4.30) that $|X(j\omega)|$ is an even function of ω and $\sphericalangle X(j\omega)$ is an odd function of ω. Thus, when computing or displaying the Fourier transform of a real-valued signal, the real and imaginary parts or magnitude and phase of the transform need only be specified for positive frequencies, as the values for negative frequencies can be determined directly from the values for $\omega > 0$ using the relationships just derived.

As a further consequence of eq. (4.30), if $x(t)$ is both real and even, then $X(j\omega)$ will also be real and even. To see this, we write

$$X(-j\omega) = \int_{-\infty}^{+\infty} x(t)e^{j\omega t}\,dt,$$

or, with the substitution $\tau = -t$,

$$X(-j\omega) = \int_{-\infty}^{+\infty} x(-\tau)e^{-j\omega\tau}d\tau.$$

Since $x(-\tau) = x(\tau)$, we have

$$\begin{aligned} X(-j\omega) &= \int_{-\infty}^{+\infty} x(\tau)e^{-j\omega t}d\tau \\ &= X(j\omega). \end{aligned}$$

Thus, $X(j\omega)$ is an even function. This, together with eq. (4.30), also requires that $X^*(j\omega) = X(j\omega)$ [i.e., that $X(j\omega)$ is real]. Example 4.2 illustrates this property for the real, even signal $e^{-a|t|}$. In a similar manner, it can be shown that if $x(t)$ is a real and odd function of time, so that $x(t) = -x(-t)$, then $X(j\omega)$ is purely imaginary and odd.

Finally, as was discussed in Chapter 1, a real function $x(t)$ can always be expressed in terms of the sum of an even function $x_e(t) = \mathcal{E}v\{x(t)\}$ and an odd function $x_o(t) = \mathcal{O}d\{x(t)\}$; that is,

$$x(t) = x_e(t) + x_o(t).$$

From the linearity of the Fourier transform,

$$\mathfrak{F}\{x(t)\} = \mathfrak{F}\{x_e(t)\} + \mathfrak{F}\{x_o(t)\},$$

and from the preceding discussion, $\mathfrak{F}\{x_e(t)\}$ is a real function and $\mathfrak{F}\{x_o(t)\}$ is purely imaginary. Thus, we can conclude that, with $x(t)$ real,

$$x(t) \overset{\mathfrak{F}}{\longleftrightarrow} X(j\omega),$$

$$\mathcal{E}v\{x(t)\} \overset{\mathfrak{F}}{\longleftrightarrow} \mathcal{R}e\{X(j\omega)\},$$

$$\mathcal{O}d\{x(t)\} \overset{\mathfrak{F}}{\longleftrightarrow} j\mathcal{I}m\{X(j\omega)\}.$$

One use of these symmetry properties is illustrated in the following example.

Example 4.10

Consider again the Fourier transform evaluation of Example 4.2 for the signal $x(t) = e^{-a|t|}$, where $a > 0$. This time we will utilize the symmetry properties of the Fourier transform to aid the evaluation process.

From Example 4.1, we have

$$e^{-at}u(t) \overset{\mathfrak{F}}{\longleftrightarrow} \frac{1}{a + j\omega}.$$

Note that for $t > 0$, $x(t)$ equals $e^{-at}u(t)$, while for $t < 0$, $x(t)$ takes on mirror image values. That is,

$$\begin{aligned} x(t) &= e^{-a|t|} = e^{-at}u(t) + e^{at}u(-t) \\ &= 2\left[\frac{e^{-at}u(t) + e^{at}u(-t)}{2}\right] \\ &= 2\mathcal{E}v\{e^{-at}u(t)\}. \end{aligned}$$

Since $e^{-at}u(t)$ is real valued, the symmetry properties of the Fourier transform lead us to conclude that

$$\mathcal{E}v\{e^{-at}u(t)\} \overset{\mathfrak{F}}{\longleftrightarrow} \mathcal{R}e\left\{\frac{1}{a+j\omega}\right\}.$$

It follows that

$$X(j\omega) = 2\mathcal{R}e\left\{\frac{1}{a+j\omega}\right\} = \frac{2a}{a^2+\omega^2},$$

which is the same as the answer found in Example 4.2.

4.3.4 Differentiation and Integration

Let $x(t)$ be a signal with Fourier transform $X(j\omega)$. Then, by differentiating both sides of the Fourier transform synthesis equation (4.24), we obtain

$$\frac{dx(t)}{dt} = \frac{1}{2\pi}\int_{-\infty}^{+\infty} j\omega X(j\omega)e^{j\omega t}d\omega.$$

Therefore,

$$\boxed{\frac{dx(t)}{dt} \overset{\mathfrak{F}}{\longleftrightarrow} j\omega X(j\omega).} \tag{4.31}$$

This is a particularly important property, as it replaces the operation of differentiation in the time domain with that of multiplication by $j\omega$ in the frequency domain. We will find the substitution to be extremely useful in our discussion in Section 4.7 on the use of Fourier transforms for the analysis of LTI systems described by differential equations.

Since differentiation in the time domain corresponds to multiplication by $j\omega$ in the frequency domain, one might conclude that integration should involve division by $j\omega$ in the frequency domain. This is indeed the case, but it is only one part of the picture. The precise relationship is

$$\boxed{\int_{-\infty}^{t} x(\tau)d\tau \overset{\mathfrak{F}}{\longleftrightarrow} \frac{1}{j\omega}X(j\omega) + \pi X(0)\delta(\omega).} \tag{4.32}$$

The impulse term on the right-hand side of eq. (4.32) reflects the dc or average value that can result from integration.

The use of eqs. (4.31) and (4.32) is illustrated in the next two examples.

Example 4.11

Let us determine the Fourier transform $X(j\omega)$ of the unit step $x(t) = u(t)$, making use of eq. (4.32) and the knowledge that

$$g(t) = \delta(t) \overset{\mathfrak{F}}{\longleftrightarrow} G(j\omega) = 1.$$

Noting that

$$x(t) = \int_{-\infty}^{t} g(\tau)d\tau$$

and taking the Fourier transform of both sides, we obtain

$$X(j\omega) = \frac{G(j\omega)}{j\omega} + \pi G(0)\delta(\omega),$$

where we have used the integration property listed in Table 4.1. Since $G(j\omega) = 1$, we conclude that

$$X(j\omega) = \frac{1}{j\omega} + \pi\delta(\omega). \tag{4.33}$$

Observe that we can apply the differentiation property of eq. (4.31) to recover the transform of the impulse. That is,

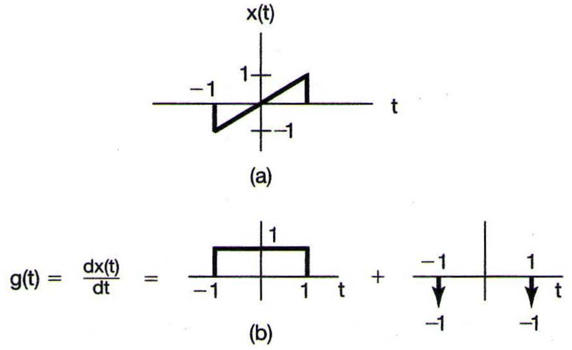
$$\delta(t) = \frac{du(t)}{dt} \overset{\mathfrak{F}}{\longleftrightarrow} j\omega\left[\frac{1}{j\omega} + \pi\delta(\omega)\right] = 1,$$

where the last equality follows from the fact that $\omega\delta(\omega) = 0$.

Example 4.12

Suppose that we wish to calculate the Fourier transform $X(j\omega)$ for the signal $x(t)$ displayed in Figure 4.16(a). Rather than applying the Fourier integral directly to $x(t)$, we instead consider the signal

$$g(t) = \frac{d}{dt}x(t).$$



(a)

(b)

Figure 4.16 (a) A signal $x(t)$ for which the Fourier transform is to be evaluated; (b) representation of the derivative of $x(t)$ as the sum of two components.

As illustrated in Figure 4.16(b), $g(t)$ is the sum of a rectangular pulse and two impulses. The Fourier transforms of each of these component signals may be determined from Table 4.2:

$$G(j\omega) = \left(\frac{2\sin\omega}{\omega}\right) - e^{j\omega} - e^{-j\omega}.$$

Note that $G(0) = 0$. Using the integration property, we obtain

$$X(j\omega) = \frac{G(j\omega)}{j\omega} + \pi G(0)\delta(\omega).$$

With $G(0) = 0$ this becomes

$$X(j\omega) = \frac{2\sin\omega}{j\omega^2} - \frac{2\cos\omega}{j\omega}.$$

The expression for $X(j\omega)$ is purely imaginary and odd, which is consistent with the fact that $x(t)$ is real and odd.

4.3.5 Time and Frequency Scaling

If

$$x(t) \overset{\mathfrak{F}}{\longleftrightarrow} X(j\omega),$$

then

$$\boxed{x(at) \overset{\mathfrak{F}}{\longleftrightarrow} \frac{1}{|a|}X\left(\frac{j\omega}{a}\right),} \tag{4.34}$$

where a is a nonzero real number. This property follows directly from the definition of the Fourier transform—specifically,

$$\mathfrak{F}\{x(at)\} = \int_{-\infty}^{+\infty} x(at)e^{-j\omega t}dt.$$

Using the substitution $\tau = at$, we obtain

$$\mathfrak{F}\{x(at)\} = \begin{cases} \dfrac{1}{a}\displaystyle\int_{-\infty}^{+\infty} x(\tau)e^{-j(\omega/a)\tau}d\tau, & a > 0 \\ -\dfrac{1}{a}\displaystyle\int_{-\infty}^{+\infty} x(\tau)e^{-j(\omega/a)\tau}d\tau, & a < 0 \end{cases},$$

which corresponds to eq. (4.34). Thus, aside from the amplitude factor $1/|a|$, a linear scaling in time by a factor of a corresponds to a linear scaling in frequency by a factor of $1/a$, and vice versa. Also, letting $a = -1$, we see from eq. (4.34) that

$$\boxed{x(-t) \overset{\mathcal{F}}{\longleftrightarrow} X(-j\omega).} \tag{4.35}$$

That is, reversing a signal in time also reverses its Fourier transform.

A common illustration of eq. (4.34) is the effect on frequency content that results when an audiotape is recorded at one speed and played back at a different speed. If the playback speed is higher than the recording speed, corresponding to compression in time (i.e., $a > 1$), then the spectrum is expanded in frequency (i.e., the audible effect is that the playback frequencies are higher). Conversely, the signal played back will be scaled down in frequency if the playback speed is slower than the recording speed ($0 < a < 1$). For example, if a recording of the sound of a small bell ringing is played back at a reduced speed, the result will sound like the chiming of a larger and deeper sounding bell.

The scaling property is another example of the inverse relationship between time and frequency that we have already encountered on several occasions. For example, we have seen that as we increase the period of a sinusoidal signal, we decrease its frequency. Also, as we saw in Example 4.5 (see Figure 4.11), if we consider the transform

$$X(j\omega) = \begin{cases} 1, & |\omega| < W \\ 0, & |\omega| > W \end{cases},$$

then as we increase W, the inverse transform of $X(j\omega)$ becomes narrower and taller and approaches an impulse as $W \to \infty$. Finally, in Example 4.8, we saw that the spacing in the frequency domain between impulses in the Fourier transform of a periodic impulse train is inversely proportional to the spacing in the time domain.

The inverse relationship between the time and frequency domains is of great importance in a variety of signal and systems contexts, including filtering and filter design, and we will encounter its consequences on numerous occasions in the remainder of the book. In addition, the reader may very well come across the implications of this property in studying a wide variety of other topics in science and engineering. One example is the uncertainty principle in physics; another is illustrated in Problem 4.49.

4.3.6 Duality

By comparing the transform and inverse transform relations given in eqs. (4.24) and (4.25), we observe that these equations are similar, but not quite identical, in form. This symmetry leads to a property of the Fourier transform referred to as *duality.* In Example 4.5, we alluded to duality when we noted the relationship that exists between the Fourier transform pairs of Examples 4.4 and 4.5. In the former example we derived the Fourier transform pair

$$x_1(t) = \begin{cases} 1, & |t| < T_1 \\ 0, & |t| > T_1 \end{cases} \overset{\mathcal{F}}{\longleftrightarrow} X_1(j\omega) = \frac{2\sin\omega T_1}{\omega}, \tag{4.36}$$

while in the latter we considered the pair

$$x_2(t) = \frac{\sin Wt}{\pi t} \overset{\mathcal{F}}{\longleftrightarrow} X_2(j\omega) = \begin{cases} 1, & |\omega| < W \\ 0, & |\omega| > W \end{cases}. \tag{4.37}$$

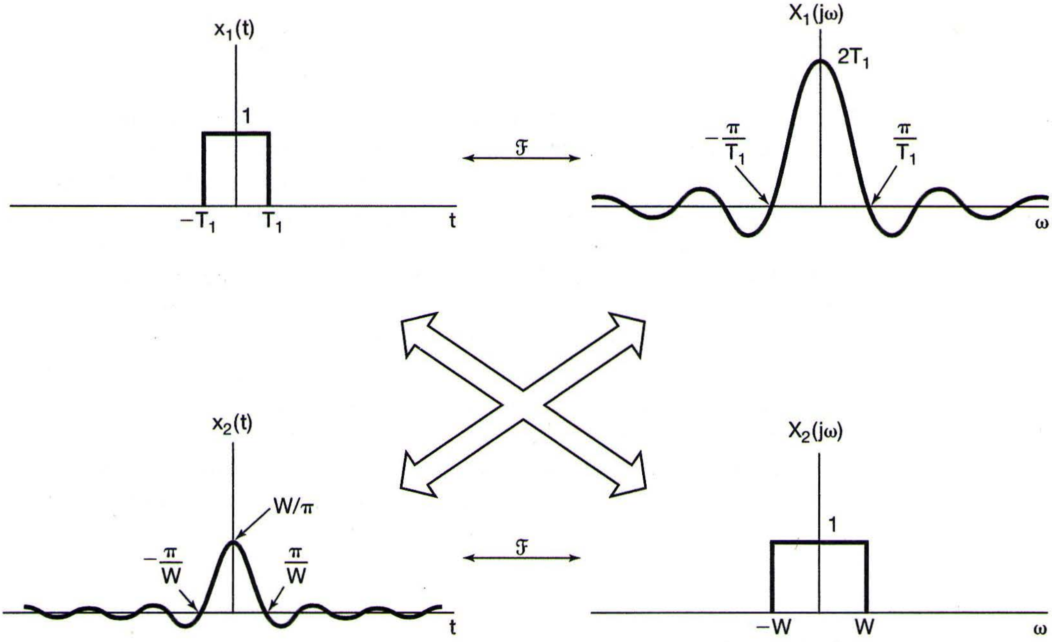

$x_1(t)$
1
$-T_1$
T_1
t
$\mathfrak{F}$
$X_1(j\omega)$
$2T_1$
$-\frac{\pi}{T_1}$
$\frac{\pi}{T_1}$
ω
$x_2(t)$
W/π
$-\frac{\pi}{W}$
$\frac{\pi}{W}$
t
$\mathfrak{F}$
$X_2(j\omega)$
1
$-W$
W
ω

Figure 4.17 Relationship between the Fourier transform pairs of eqs. (4.36) and (4.37).

The two Fourier transform pairs and the relationship between them are depicted in Figure 4.17.

The symmetry exhibited by these two examples extends to Fourier transforms in general. Specifically, because of the symmetry between eqs. (4.24) and (4.25), for any transform pair, there is a dual pair with the time and frequency variables interchanged. This is best illustrated through an example.

Example 4.13

Let us consider using duality to find the Fourier transform $G(j\omega)$ of the signal

$$g(t) = \frac{2}{1+t^2}.$$

In Example 4.2 we encountered a Fourier transform pair in which the Fourier transform, as a function of ω, had a form similar to that of the signal $x(t)$. Specifically, suppose we consider a signal $x(t)$ whose Fourier transform is

$$X(j\omega) = \frac{2}{1+\omega^2}.$$

Then, from Example 4.2,

$$x(t) = e^{-|t|} \overset{\mathfrak{F}}{\longleftrightarrow} X(j\omega) = \frac{2}{1+\omega^2}.$$

The synthesis equation for this Fourier transform pair is

$$e^{-|t|} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\frac{2}{1+\omega^2}\right)e^{j\omega t}d\omega.$$

Multiplying this equation by 2π and replacing t by $-t$, we obtain

$$2\pi e^{-|t|} = \int_{-\infty}^{\infty}\left(\frac{2}{1+\omega^2}\right)e^{-j\omega t}d\omega.$$

Now, interchanging the names of the variables t and ω, we find that

$$2\pi e^{-|\omega|} = \int_{-\infty}^{\infty}\left(\frac{2}{1+t^2}\right)e^{-j\omega t}dt. \tag{4.38}$$

The right-hand side of eq. (4.38) is the Fourier transform analysis equation for $2/(1+t^2)$, and thus, we conclude that

$$\mathfrak{F}\left\{\frac{2}{1+t^2}\right\} = 2\pi e^{-|\omega|}.$$

The duality property can also be used to determine or to suggest other properties of Fourier transforms. Specifically, if there are characteristics of a function of time that have implications with regard to the Fourier transform, then the same characteristics associated with a function of frequency will have *dual* implications in the time domain. For example, in Section 4.3.4, we saw that differentiation in the time domain corresponds to multiplication by $j\omega$ in the frequency domain. From the preceding discussion, we might then suspect that multiplication by jt in the time domain corresponds roughly to differentiation in the frequency domain. To determine the precise form of this dual property, we can proceed in a fashion exactly analogous to that used in Section 4.3.4. Thus, if we differentiate the analysis equation (4.25) with respect to ω, we obtain

$$\frac{dX(j\omega)}{d\omega} = \int_{-\infty}^{+\infty} -jtx(t)e^{-j\omega t}dt. \tag{4.39}$$

That is,

$$\boxed{-jtx(t) \overset{\mathfrak{F}}{\longleftrightarrow} \frac{dX(j\omega)}{d\omega}.} \tag{4.40}$$

Similarly, we can derive the dual properties of eqs. (4.27) and (4.32):

$$\boxed{e^{j\omega_0 t}x(t) \overset{\mathfrak{F}}{\longleftrightarrow} X(j(\omega-\omega_0))} \tag{4.41}$$

and

$$\boxed{-\frac{1}{jt}x(t) + \pi x(0)\delta(t) \overset{\mathfrak{F}}{\longleftrightarrow} \int_{-\infty}^{\omega} x(\eta)d\eta.} \tag{4.42}$$

4.3.7 Parseval's Relation

If $x(t)$ and $X(j\omega)$ are a Fourier transform pair, then

$$\boxed{\int_{-\infty}^{+\infty}|x(t)|^2dt = \frac{1}{2\pi}\int_{-\infty}^{+\infty}|X(j\omega)|^2d\omega.} \tag{4.43}$$

This expression, referred to as Parseval's relation, follows from direct application of the Fourier transform. Specifically,

$$\begin{aligned}\int_{-\infty}^{+\infty}|x(t)|^2dt &= \int_{-\infty}^{+\infty}x(t)x^*(t)dt\\ &= \int_{-\infty}^{+\infty}x(t)\left[\frac{1}{2\pi}\int_{-\infty}^{+\infty}X^*(j\omega)e^{-j\omega t}d\omega\right]dt.\end{aligned}$$

Reversing the order of integration gives

$$\int_{-\infty}^{+\infty}|x(t)|^2dt = \frac{1}{2\pi}\int_{-\infty}^{+\infty}X^*(j\omega)\left[\int_{-\infty}^{+\infty}x(t)e^{-j\omega t}dt\right]d\omega.$$

The bracketed term is simply the Fourier transform of $x(t)$; thus,

$$\int_{-\infty}^{+\infty}|x(t)|^2dt = \frac{1}{2\pi}\int_{-\infty}^{+\infty}|X(j\omega)|^2d\omega.$$

The term on the left-hand side of eq. (4.43) is the total energy in the signal $x(t)$. Parseval's relation says that this total energy may be determined either by computing the energy per unit time ($|x(t)|^2$) and integrating over all time or by computing the energy per unit frequency ($|X(j\omega)|^2/2\pi$) and integrating over all frequencies. For this reason, $|X(j\omega)|^2$ is often referred to as the *energy-density spectrum* of the signal $x(t)$. (See also Problem 4.45.) Note that Parseval's relation for finite-energy signals is the direct counterpart of Parseval's relation for periodic signals (eq. 3.67), which states that the average *power* of a periodic signal equals the sum of the average powers of its individual harmonic components, which in turn are equal to the squared magnitudes of the Fourier series coefficients.

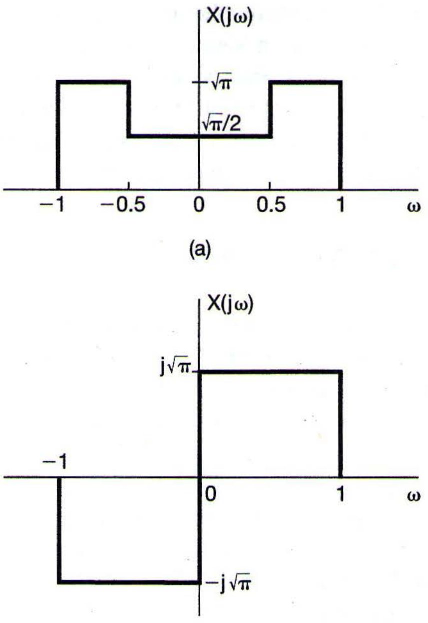
Parseval's relation and other Fourier transform properties are often useful in determining some time domain characteristics of a signal directly from the Fourier transform. The next example is a simple illustration of this.

Example 4.14

For each of the Fourier transforms shown in Figure 4.18, we wish to evaluate the following time-domain expressions:

$$E = \int_{-\infty}^{\infty}|x(t)|^2dt$$

$$D = \left.\frac{d}{dt}x(t)\right|_{t=0}$$



(a)

(b)

Figure 4.18 The Fourier transforms considered in Example 4.14.

To evaluate E in the frequency domain, we may use Parseval's relation. That is,

$$E = \frac{1}{2\pi}\int_{-\infty}^{\infty} |X(j\omega)|^2 d\omega \tag{4.44}$$

which evaluates to $\frac{5}{8}$ for Figure 4.18(a) and to 1 for Figure 4.18(b).

To evaluate D in the frequency domain, we first use the differentiation property to observe that

$$g(t) = \frac{d}{dt}x(t) \overset{\mathfrak{F}}{\longleftrightarrow} j\omega X(j\omega) = G(j\omega).$$

Noting that

$$D = g(0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} G(j\omega) d\omega \tag{4.45}$$

we conclude:

$$D = \int_{-\infty}^{\infty} j\omega X(j\omega)\, d\omega \tag{4.46}$$

which evaluates to zero for figure 4.18(a) and to $\dfrac{-1}{(2\sqrt{\pi})}$ for Figure 4.18(b).

There are many other properties of the Fourier transform in addition to those we have already discussed. In the next two sections, we present two specific properties that play particularly central roles in the study of LTI systems and their applications. The first of these, discussed in Section 4.4, is referred to as the *convolution property,* which is central to many signals and systems applications, including filtering. The second, discussed in Section 4.5, is referred to as the *multiplication property,* and it provides the foundation for our discussion of sampling in Chapter 7 and amplitude modulation in Chapter 8. In Section 4.6, we summarize the properties of the Fourier transform.

4.4 THE CONVOLUTION PROPERTY

As we saw in Chapter 3, if a periodic signal is represented in a Fourier series—i.e., as a linear combination of harmonically related complex exponentials, as in eq. (3.38)—then the response of an LTI system to this input can also be represented by a Fourier series. Because complex exponentials are eigenfunctions of LTI systems, the Fourier series coefficients of the output are those of the input multiplied by the frequency response of the system evaluated at the corresponding harmonic frequencies.

In this section, we extend this result to the situation in which the signals are aperiodic. We first derive the property somewhat informally, to build on the intuition we developed for periodic signals in Chapter 3, and then provide a brief, formal derivation starting directly from the convolution integral.

Recall our interpretation of the Fourier transform synthesis equation as an expression for $x(t)$ as a linear combination of complex exponentials. Specifically, referring back to eq. (4.7), $x(t)$ is expressed as the limit of a sum; that is,

$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} X(j\omega)e^{j\omega t}d\omega = \lim_{\omega_0\to 0}\frac{1}{2\pi}\sum_{k=-\infty}^{+\infty} X(jk\omega_0)e^{jk\omega_0 t}\omega_0. \tag{4.47}$$

As developed in Sections 3.2 and 3.8, the response of a linear system with impulse response $h(t)$ to a complex exponential $e^{jk\omega_0 t}$ is $H(jk\omega_0)e^{jk\omega_0 t}$, where

$$H(jk\omega_0) = \int_{-\infty}^{+\infty} h(t)e^{-jk\omega_0 t}dt. \tag{4.48}$$

We can recognize the frequency response $H(j\omega)$, as defined in eq. (3.121), as the Fourier transform of the system impulse response. In other words, the Fourier transform of the impulse response (evaluated at $\omega = k\omega_0$) is the complex scaling factor that the LTI system applies to the eigenfunction $e^{jk\omega_0 t}$. From superposition [see eq. (3.124)], we then have

$$\frac{1}{2\pi}\sum_{k=-\infty}^{+\infty} X(jk\omega_0)e^{jk\omega_0 t}\omega_0 \longrightarrow \frac{1}{2\pi}\sum_{k=-\infty}^{+\infty} X(jk\omega_0)H(jk\omega_0)e^{jk\omega_0 t}\omega_0,$$

and thus, from eq. (4.47), the response of the linear system to $x(t)$ is

$$\begin{aligned} y(t) &= \lim_{\omega_0\to 0}\frac{1}{2\pi}\sum_{k=-\infty}^{+\infty} X(jk\omega_0)H(jk\omega_0)e^{jk\omega_0 t}\omega_0 \\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} X(j\omega)H(j\omega)e^{j\omega t}d\omega. \end{aligned} \tag{4.49}$$

Since $y(t)$ and its Fourier transform $Y(j\omega)$ are related by

$$y(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} Y(j\omega)e^{j\omega t}d\omega, \tag{4.50}$$

we can identify $Y(j\omega)$ from eq. (4.49), yielding

$$Y(j\omega) = X(j\omega)H(j\omega). \tag{4.51}$$

As a more formal derivation, we consider the convolution integral

$$y(t) = \int_{-\infty}^{+\infty} x(\tau)h(t-\tau)d\tau. \tag{4.52}$$

We desire $Y(j\omega)$, which is

$$Y(j\omega) = \mathfrak{F}\{y(t)\} = \int_{-\infty}^{+\infty}\left[\int_{-\infty}^{+\infty} x(\tau)h(t-\tau)d\tau\right]e^{-j\omega t}dt. \tag{4.53}$$

Interchanging the order of integration and noting that $x(\tau)$ does not depend on t, we have

$$Y(j\omega) = \int_{-\infty}^{+\infty} x(\tau)\left[\int_{-\infty}^{+\infty} h(t-\tau)e^{-j\omega t}dt\right]d\tau. \tag{4.54}$$

By the time-shift property, eq. (4.27), the bracketed term is $e^{-j\omega\tau}H(j\omega)$. Substituting this into eq. (4.54) yields

$$Y(j\omega) = \int_{-\infty}^{+\infty} x(\tau)e^{-j\omega\tau}H(j\omega)d\tau = H(j\omega)\int_{-\infty}^{+\infty} x(\tau)e^{-j\omega\tau}d\tau. \tag{4.55}$$

The integral is $X(j\omega)$, and hence,

$$Y(j\omega) = H(j\omega)X(j\omega).$$

That is,

$$\boxed{y(t) = h(t) * x(t) \overset{\mathfrak{F}}{\longleftrightarrow} Y(j\omega) = H(j\omega)X(j\omega).} \tag{4.56}$$

Equation (4.56) is of major importance in signal and system analysis. As expressed in this equation, the Fourier transform maps the convolution of two signals into the product of their Fourier transforms. $H(j\omega)$, the Fourier transform of the impulse response, is the frequency response as defined in eq. (3.121) and captures the change in complex amplitude of the Fourier transform of the input at each frequency ω. For example, in frequency-selective filtering we may want to have $H(j\omega) \approx 1$ over one range of frequencies, so that the frequency components in this band experience little or no attenuation or change due to the system, while over another range of frequencies we may want to have $H(j\omega) \approx 0$, so that components in this range are eliminated or significantly attenuated.

The frequency response $H(j\omega)$ plays as important a role in the analysis of LTI systems as does its inverse transform, the unit impulse response. For one thing, since $h(t)$ completely characterizes an LTI system, then so must $H(j\omega)$. In addition, many of the properties of LTI systems can be conveniently interpreted in terms of $H(j\omega)$. For example, in Section 2.3, we saw that the impulse response of the cascade of two LTI systems is the convolution of the impulse responses of the individual systems and that the overall impulse response does not depend on the order in which the systems are cascaded. Using eq. (4.56), we can rephrase this in terms of frequency responses. As illustrated in Figure 4.19, since the impulse response of the cascade of two LTI systems is the convolution of the individual impulse responses, the convolution property then implies that the overall frequency response of the cascade of two systems is simply the product of the individual frequency responses. From this observation, it is then clear that the overall frequency response does not depend on the order of the cascade.

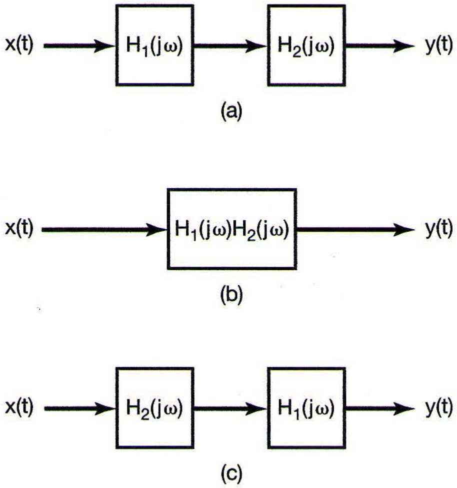


Figure 4.19 Three equivalent LTI systems. Here, each block represents an LTI system with the indicated frequency response.

As discussed in Section 4.1.2, convergence of the Fourier transform is guaranteed only under certain conditions, and consequently, the frequency response cannot be defined for every LTI system. If, however, an LTI system is stable, then, as we saw in Section 2.3.7 and Problem 2.49, its impulse response is absolutely integrable; that is,

$$\int_{-\infty}^{+\infty} |h(t)|dt < \infty. \tag{4.57}$$

Equation (4.57) is one of the three Dirichlet conditions that together guarantee the existence of the Fourier transform $H(j\omega)$ of $h(t)$. Thus, assuming that $h(t)$ satisfies the other two conditions, as essentially all signals of physical or practical significance do, we see that a stable LTI system has a frequency response $H(j\omega)$.

In using Fourier analysis to study LTI systems, we will be restricting ourselves to systems whose impulse responses possess Fourier transforms. In order to use transform techniques to examine unstable LTI systems we will develop a generalization of the continuous-time Fourier transform, the Laplace transform. We defer this discussion to Chapter 9, and until then we will consider the many problems and practical applications that we can analyze using the Fourier transform.

4.4.1 Examples

To illustrate the convolution property and its applications further, let us consider several examples.

Example 4.15

Consider a continuous-time LTI system with impulse response

$$h(t) = \delta(t - t_0). \tag{4.58}$$

The frequency response of this system is the Fourier transform of $h(t)$ and is given by

$$H(j\omega) = e^{-j\omega t_0}. \tag{4.59}$$

Thus, for any input $x(t)$ with Fourier transform $X(j\omega)$, the Fourier transform of the output is

$$\begin{aligned} Y(j\omega) &= H(j\omega)X(j\omega) \\ &= e^{-j\omega t_0}X(j\omega). \end{aligned} \tag{4.60}$$

This result, in fact, is consistent with the time-shift property of Section 4.3.2. Specifically, a system for which the impulse response is $\delta(t - t_0)$ applies a time shift of t_0 to the input—that is,

$$y(t) = x(t - t_0).$$

Thus, the shifting property given in eq. (4.27) also yields eq. (4.60). Note that, either from our discussion in Section 4.3.2 or directly from eq. (4.59), the frequency response of a system that is a pure time shift has unity magnitude at all frequencies (i.e., $|e^{-j\omega t_0}| = 1$) and has a phase characteristic $-\omega t_0$ that is a linear function of ω.

Example 4.16

As a second example, let us examine a differentiator—that is, an LTI system for which the input $x(t)$ and the output $y(t)$ are related by

$$y(t) = \frac{dx(t)}{dt}.$$

From the differentiation property of Section 4.3.4,

$$Y(j\omega) = j\omega X(j\omega). \tag{4.61}$$

Consequently, from eq. (4.56), it follows that the frequency response of a differentiator is

$$H(j\omega) = j\omega. \tag{4.62}$$

Example 4.17

Consider an integrator—that is, an LTI system specified by the equation

$$y(t) = \int_{-\infty}^{t} x(\tau)d\tau.$$

The impulse response for this system is the unit step $u(t)$, and therefore, from Example 4.11 and eq. (4.33), the frequency response of the system is

$$H(j\omega) = \frac{1}{j\omega} + \pi\delta(\omega).$$

Then using eq. (4.56), we have

$$\begin{aligned} Y(j\omega) &= H(j\omega)X(j\omega) \\ &= \frac{1}{j\omega}X(j\omega) + \pi X(j\omega)\delta(\omega) \\ &= \frac{1}{j\omega}X(j\omega) + \pi X(0)\delta(\omega), \end{aligned}$$

which is consistent with the integration property of eq. (4.32).

Example 4.18

As we discussed in Section 3.9.2, frequency-selective filtering is accomplished with an LTI system whose frequency response $H(j\omega)$ passes the desired range of frequencies and significantly attenuates frequencies outside that range. For example, consider the ideal lowpass filter introduced in Section 3.9.2, which has the frequency reponse illustrated in Figure 4.20 and given by

$$H(j\omega) = \begin{cases} 1 & |\omega| < \omega_c \\ 0 & |\omega| > \omega_c \end{cases}. \tag{4.63}$$

Now that we have developed the Fourier transform representation, we know that the impulse response $h(t)$ of this ideal filter is the inverse transform of eq. (4.63). Using the result in Example 4.5, we then have

$$h(t) = \frac{\sin \omega_c t}{\pi t}, \tag{4.64}$$

which is plotted in Figure 4.21.

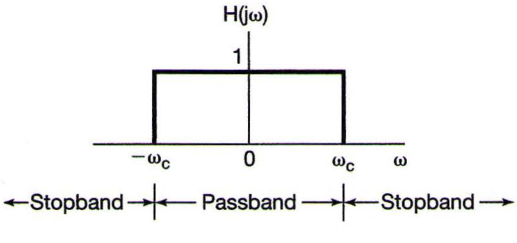


Figure 4.20 Frequency response of an ideal lowpass filter.

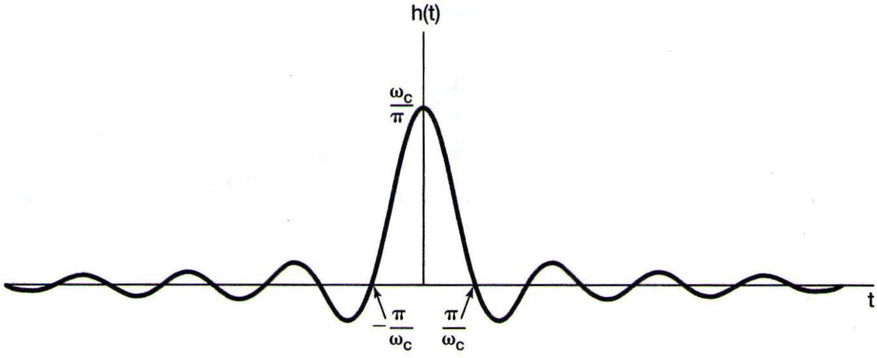


Figure 4.21 Impulse response of an ideal lowpass filter.

From Example 4.18, we can begin to see some of the issues that arise in filter design that involve looking in both the time and frequency domains. In particular, while the ideal lowpass filter does have perfect frequency selectivity, its impulse response has some characteristics that may not be desirable. First, note that $h(t)$ is not zero for $t < 0$. Consequently, the ideal lowpass filter is not causal, and thus, in applications requiring causal systems, the ideal filter is not an option. Moreover, as we discuss in Chapter 6, even if causality is not an essential constraint, the ideal filter is not easy to approximate closely, and nonideal filters that are more easily implemented are typically preferred. Furthermore, in some applications (such as the automobile suspension system discussed in Section 6.7.1), oscillatory behavior in the impulse response of a lowpass filter may be undesirable. In such applications the time domain characteristics of the ideal lowpass filter, as shown in Figure 4.21, may be unacceptable, implying that we may need to trade off frequency-domain characteristics such as ideal frequency selectivity with time-domain properties.

For example, consider the LTI system with impulse response

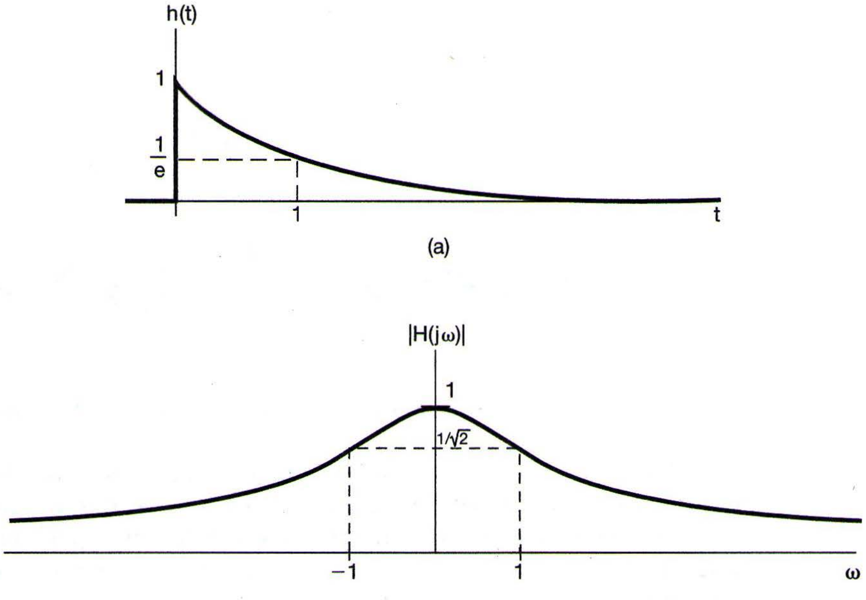
$$h(t) = e^{-t}u(t). \tag{4.65}$$

The frequency response of this system is

$$H(j\omega) = \frac{1}{j\omega + 1}. \tag{4.66}$$

Comparing eqs. (3.145) and (4.66), we see that this system can be implemented with the simple *RC* circuit discussed in Section 3.10. The impulse response and the magnitude of the frequency response are shown in Figure 4.22. While the system does not have the strong frequency selectivity of the ideal lowpass filter, it is causal and has an impulse response that decays monotonically, i.e., without oscillations. This filter or somewhat more complex ones corresponding to higher order differential equations are quite frequently preferred to ideal filters because of their causality, ease of implementation, and flexibility in allowing trade-offs, among other design considerations such as frequency selectivity and oscillatory behavior in the time domain. Many of these issues will be discussed in more detail in Chapter 6.

The convolution property is often useful in evaluating the convolution integral—i.e., in computing the response of LTI systems. This is illustrated in the next example.



(a)

(b)

Figure 4.22 (a) Impulse response of the LTI system in eq. (4.65); (b) magnitude of the frequency response of the system.

Example 4.19

Consider the response of an LTI system with impulse response

$$h(t) = e^{-at}u(t), \quad a > 0,$$

to the input signal

$$x(t) = e^{-bt}u(t), \quad b > 0.$$

Rather than computing $y(t) = x(t) * h(t)$ directly, let us transform the problem into the frequency domain. From Example 4.1, the Fourier transforms of $x(t)$ and $h(t)$ are

$$X(j\omega) = \frac{1}{b + j\omega}$$

and

$$H(j\omega) = \frac{1}{a + j\omega}.$$

Therefore,

$$Y(j\omega) = \frac{1}{(a + j\omega)(b + j\omega)}. \tag{4.67}$$

To determine the output $y(t)$, we wish to obtain the inverse transform of $Y(j\omega)$. This is most simply done by expanding $Y(j\omega)$ in a partial-fraction expansion. Such expansions are extremely useful in evaluating inverse transforms, and the general method for performing a partial-fraction expansion is developed in the appendix. For this example, assuming that $b \neq a$, the partial fraction expansion for $Y(j\omega)$ takes the form

$$Y(j\omega) = \frac{A}{a + j\omega} + \frac{B}{b + j\omega}, \tag{4.68}$$

where A and B are constants to be determined. One way to find A and B is to equate the right-hand sides of eqs. (4.67) and (4.68), multiply both sides by $(a + j\omega)(b + j\omega)$, and solve for A and B. Alternatively, in the appendix we present a more general and efficient method for computing the coefficients in partial-fraction expansions such as eq. (4.68). Using either of these approaches, we find that

$$A = \frac{1}{b - a} = -B,$$

and therefore,

$$Y(j\omega) = \frac{1}{b - a}\left[\frac{1}{a + j\omega} - \frac{1}{b + j\omega}\right]. \tag{4.69}$$

The inverse transform for each of the two terms in eq. (4.69) can be recognized by inspection. Using the linearity property of Section 4.3.1, we have

$$y(t) = \frac{1}{b - a}[e^{-at}u(t) - e^{-bt}u(t)].$$

When $b = a$, the partial fraction expansion of eq. (4.69) is not valid. However, with $b = a$, eq. (4.67) becomes

$$Y(j\omega) = \frac{1}{(a + j\omega)^2}.$$

Recognizing this as

$$\frac{1}{(a + j\omega)^2} = j\frac{d}{d\omega}\left[\frac{1}{a + j\omega}\right],$$

we can use the dual of the differentiation property, as given in eq. (4.40). Thus,

$$e^{-at}u(t) \overset{\mathcal{F}}{\longleftrightarrow} \frac{1}{a + j\omega}$$

$$te^{-at}u(t) \overset{\mathcal{F}}{\longleftrightarrow} j\frac{d}{d\omega}\left[\frac{1}{a + j\omega}\right] = \frac{1}{(a + j\omega)^2},$$

and consequently,

$$y(t) = te^{-at}u(t).$$

Example 4.20

As another illustration of the usefulness of the convolution property, let us consider the problem of determining the response of an ideal lowpass filter to an input signal $x(t)$ that has the form of a sinc function. That is,

$$x(t) = \frac{\sin \omega_i t}{\pi t}.$$

Of course, the impulse response of the ideal lowpass filter is of a similar form, namely,

$$h(t) = \frac{\sin \omega_c t}{\pi t}.$$

The filter output $y(t)$ will therefore be the convolution of two sinc functions, which, as we now show, also turns out to be a sinc function. A particularly convenient way of deriving this result is to first observe that

$$Y(j\omega) = X(j\omega)H(j\omega),$$

where

$$X(j\omega) = \begin{cases} 1 & |\omega| \le \omega_i \\ 0 & \text{elsewhere} \end{cases}$$

and

$$H(j\omega) = \begin{cases} 1 & |\omega| \le \omega_c \\ 0 & \text{elsewhere} \end{cases}.$$

Therefore,

$$Y(j\omega) = \begin{cases} 1 & |\omega| \le \omega_0 \\ 0 & \text{elsewhere} \end{cases},$$

where ω_0 is the smaller of the two numbers ω_i and ω_c. Finally, the inverse Fourier transform of $Y(j\omega)$ is given by

$$y(t) = \begin{cases} \dfrac{\sin \omega_c t}{\pi t} & \text{if } \omega_c \le \omega_i \\ \dfrac{\sin \omega_i t}{\pi t} & \text{if } \omega_i \le \omega_c \end{cases}$$

That is, depending upon which of ω_c and ω_i is smaller, the output is equal to either $x(t)$ or $h(t)$.

4.5 THE MULTIPLICATION PROPERTY

The convolution property states that convolution in the *time* domain corresponds to multiplication in the *frequency* domain. Because of duality between the time and frequency domains, we would expect a dual property also to hold (i.e., that multiplication in the time domain corresponds to convolution in the frequency domain). Specifically,

$$\boxed{r(t) = s(t)p(t) \longleftrightarrow R(j\omega) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} S(j\theta)P(j(\omega-\theta))d\theta} \tag{4.70}$$

This can be shown by exploiting duality as discussed in Section 4.3.6, together with the convolution property, or by directly using the Fourier transform relations in a manner analogous to the procedure used in deriving the convolution property.

Multiplication of one signal by another can be thought of as using one signal to scale or *modulate* the amplitude of the other, and consequently, the multiplication of two signals is often referred to as *amplitude modulation.* For this reason, eq. (4.70) is sometimes referred to as the *modulation property.* As we shall see in Chapters 7 and 8, this property has several very important applications. To illustrate eq. (4.70), and to suggest one of the applications that we will discuss in subsequent chapters, let us consider several examples.

Example 4.21

Let $s(t)$ be a signal whose spectrum $S(j\omega)$ is depicted in Figure 4.23(a). Also, consider the signal

$$p(t) = \cos \omega_0 t.$$

Then

$$P(j\omega) = \pi\delta(\omega - \omega_0) + \pi\delta(\omega + \omega_0),$$

as sketched in Figure 4.23(b), and the spectrum $R(j\omega)$ of $r(t) = s(t)p(t)$ is obtained by

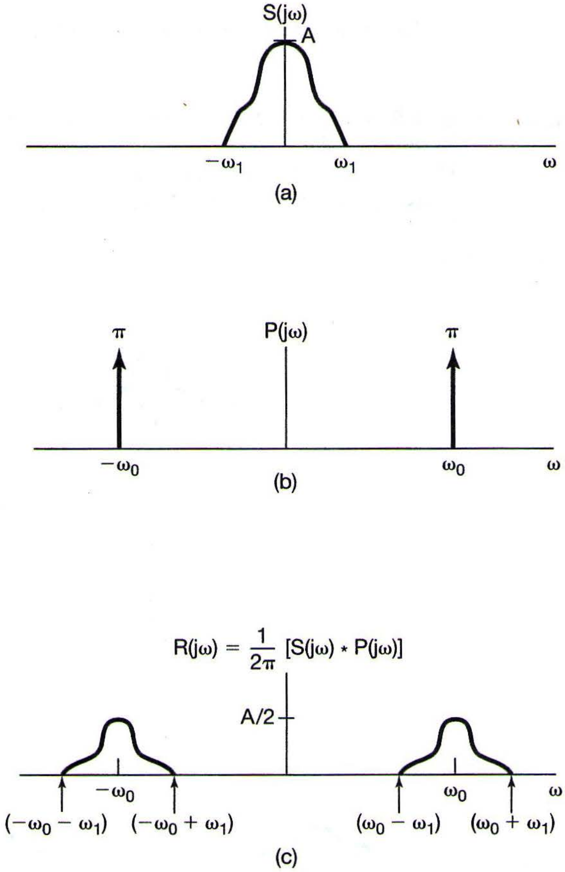


Figure 4.23 Use of the multiplication property in Example 4.21: (a) the Fourier transform of a signal $s(t)$; (b) the Fourier transform of $p(t) = \cos \omega_0 t$; (c) the Fourier transform of $r(t) = s(t)p(t)$.

an application of eq. (4.70), yielding

$$\begin{aligned} R(j\omega) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} S(j\theta)P(j(\omega-\theta))d\theta \\ &= \frac{1}{2}S(j(\omega-\omega_0)) + \frac{1}{2}S(j(\omega+\omega_0)), \end{aligned} \tag{4.71}$$

which is sketched in Figure 4.23(c). Here we have assumed that $\omega_0 > \omega_1$, so that the two nonzero portions of $R(j\omega)$ do not overlap. Clearly, the spectrum of $r(t)$ consists of the sum of two shifted and scaled versions of $S(j\omega)$.

From eq. (4.71) and from Figure 4.23, we see that all of the information in the signal $s(t)$ is preserved when we multiply this signal by a sinusoidal signal, although the information has been shifted to higher frequencies. This fact forms the basis for sinusoidal amplitude modulation systems for communications. In the next example, we learn how we can recover the original signal $s(t)$ from the amplitude-modulated signal $r(t)$.

Example 4.22

Let us now consider $r(t)$ as obtained in Example 4.21, and let

$$g(t) = r(t)p(t),$$

where, again, $p(t) = \cos\omega_0 t$. Then, $R(j\omega)$, $P(j\omega)$, and $G(j\omega)$ are as shown in Figure 4.24.

From Figure 4.24(c) and the linearity of the Fourier transform, we see that $g(t)$ is the sum of $(1/2)s(t)$ and a signal with a spectrum that is nonzero only at higher frequen-

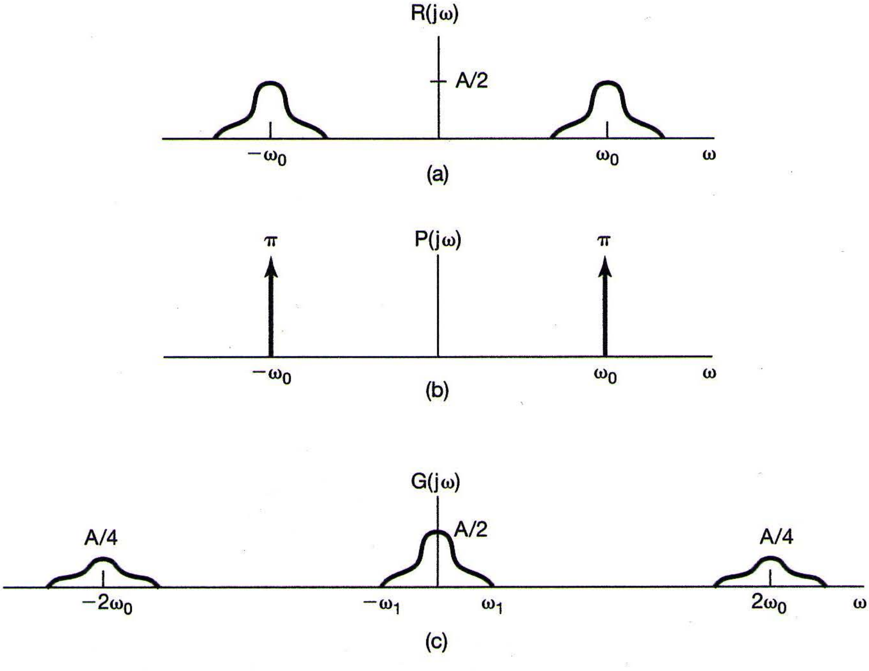


Figure 4.24 Spectra of signals considered in Example 4.22: (a) $R(j\omega)$; (b) $P(j\omega)$; (c) $G(j\omega)$.

cies (centered around $\pm 2\omega_0$). Suppose then that we apply the signal $g(t)$ as the input to a frequency-selective lowpass filter with frequency response $H(j\omega)$ that is constant at low frequencies (say, for $|\omega| < \omega_1$) and zero at high frequencies (for $|\omega| > \omega_1$). Then the output of this system will have as its spectrum $H(j\omega)G(j\omega)$, which, because of the particular choice of $H(j\omega)$, will be a scaled replica of $S(j\omega)$. Therefore, the output itself will be a scaled version of $s(t)$. In Chapter 8, we expand significantly on this idea as we develop in detail the fundamentals of amplitude modulation.

Example 4.23

Another illustration of the usefulness of the Fourier transform multiplication property is provided by the problem of determining the Fourier transform of the signal

$$x(t) = \frac{\sin(t)\sin(t/2)}{\pi t^2}.$$

The key here is to recognize $x(t)$ as the product of two sinc functions:

$$x(t) = \pi\left(\frac{\sin(t)}{\pi t}\right)\left(\frac{\sin(t/2)}{\pi t}\right).$$

Applying the multiplication property of the Fourier transform, we obtain

$$X(j\omega) = \frac{1}{2}\mathfrak{F}\left\{\frac{\sin(t)}{\pi t}\right\} * \mathfrak{F}\left\{\frac{\sin(t/2)}{\pi t}\right\}.$$

Noting that the Fourier transform of each sinc function is a rectangular pulse, we can proceed to convolve those pulses to obtain the function $X(j\omega)$ displayed in Figure 4.25.

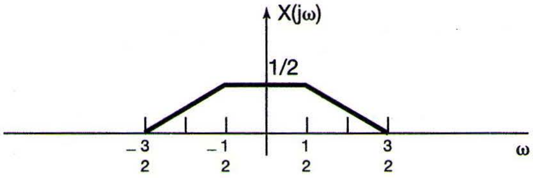
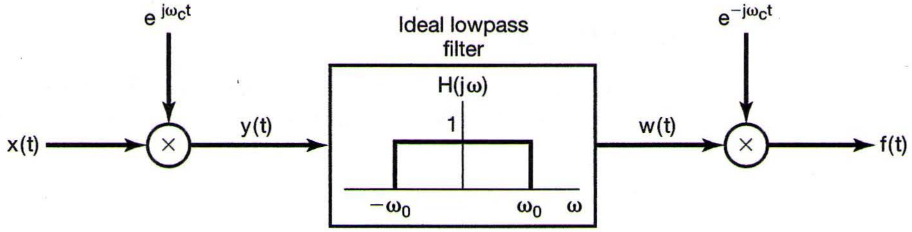


Figure 4.25 The Fourier transform of $x(t)$ in Example 4.23.

4.5.1 Frequency-Selective Filtering with Variable Center Frequency

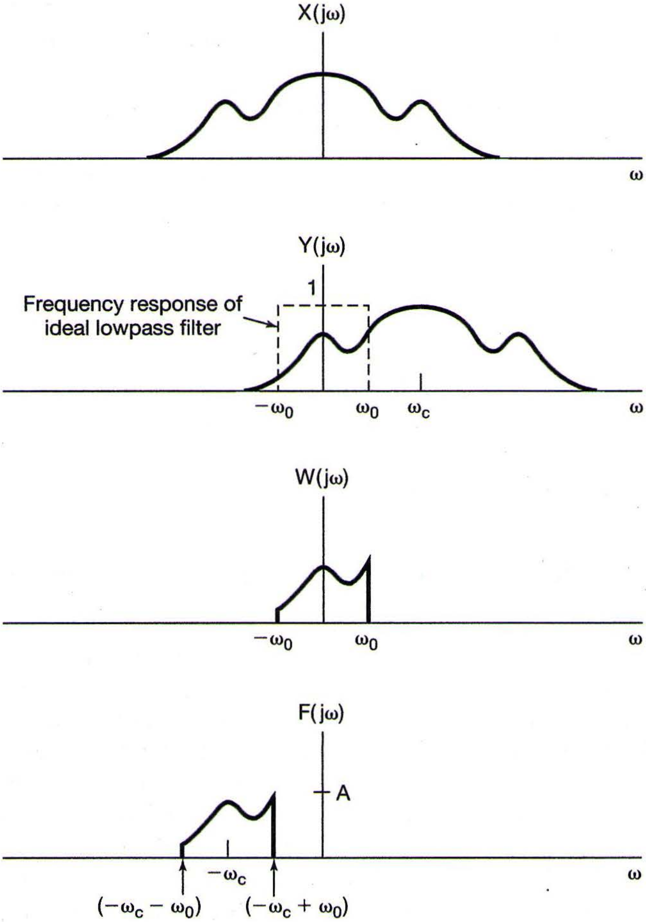
As suggested in Examples 4.21 and 4.22 and developed more fully in Chapter 8, one of the important applications of the multiplication property is amplitude modulation in communication systems. Another important application is in the implementation of frequency-selective bandpass filters with tunable center frequencies that can be adjusted by the simple turn of a dial. In a frequency-selective bandpass filter built with elements such as resistors, operational amplifiers, and capacitors, the center frequency depends on a number of element values, all of which must be varied simultaneously in the correct way if the center frequency is to be adjusted directly. This is generally difficult and cumbersome in comparison with building a filter whose characteristics are fixed. An alternative to directly varying the filter characteristics is to use a fixed frequency-selective filter and shift the spectrum of the signal appropriately, using the principles of sinusoidal amplitude modulation.

For example, consider the system shown in Figure 4.26. Here, an input signal $x(t)$ is multiplied by the complex exponential signal $e^{j\omega_c t}$. The resulting signal is then passed through a lowpass filter with cutoff frequency ω_0, and the output is multiplied by $e^{-j\omega_c t}$. The spectra of the signals $x(t)$, $y(t)$, $w(t)$, and $f(t)$ are illustrated in Figure 4.27.



$e^{j\omega_c t}$
Ideal lowpass filter
$e^{-j\omega_c t}$
H(jω)
1
x(t)
y(t)
w(t)
f(t)
$-\omega_0$
ω_0
ω

Figure 4.26 Implementation of a bandpass filter using amplitude modulation with a complex exponential carrier.



X(jω)
ω
Y(jω)
1
Frequency response of ideal lowpass filter
$-\omega_0$
ω_0
ω_c
ω
W(jω)
$-\omega_0$
ω_0
ω
F(jω)
A
$-\omega_c$
$(-\omega_c - \omega_0)$
$(-\omega_c + \omega_0)$
ω

Figure 4.27 Spectra of the signals in the system of Figure 4.26.

Specifically, from either the multiplication property or the frequency-shifting property it follows that the Fourier transform of $y(t) = e^{j\omega_c t}x(t)$ is

$$Y(j\omega) = \int_{-\infty}^{+\infty} \delta(\theta - \omega_c)X(\omega - \theta)d\theta$$

so that $Y(j\omega)$ equals $X(j\omega)$ shifted to the right by ω_c and frequencies in $X(j\omega)$ near $\omega = \omega_c$ have been shifted into the passband of the lowpass filter. Similarly, the Fourier transform of $f(t) = e^{-j\omega_c t}w(t)$ is

$$F(j\omega) = W(j(\omega + \omega_c)),$$

so that the Fourier transform of $F(j\omega)$ is $W(j\omega)$ shifted to the left by ω_c. From Figure 4.27, we observe that the overall system of Figure 4.26 is equivalent to an ideal bandpass filter with center frequency $-\omega_c$ and bandwidth $2\omega_0$, as illustrated in Figure 4.28. As the frequency ω_c of the complex exponential oscillator is varied, the center frequency of the bandpass filter varies.

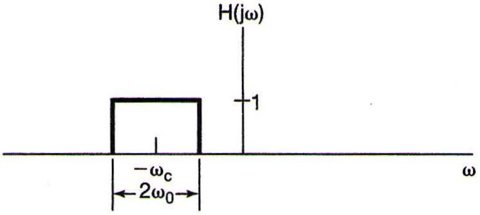


Figure 4.28 Bandpass filter equivalent of Figure 4.26.

In the system of Figure 4.26 with $x(t)$ real, the signals $y(t)$, $w(t)$, and $f(t)$ are all complex. If we retain only the real part of $f(t)$, the resulting spectrum is that shown in Figure 4.29, and the equivalent bandpass filter passes bands of frequencies centered around ω_c and $-\omega_c$, as indicated in Figure 4.30. Under certain conditions, it is also possible to use sinusoidal rather than complex exponential modulation to implement the system of the latter figure. This is explored further in Problem 4.46.

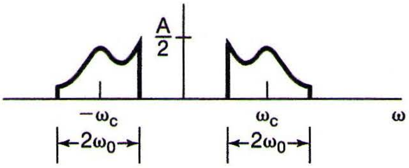


Figure 4.29 Spectrum of $\Re e\{f(t)\}$ associated with Figure 4.26.

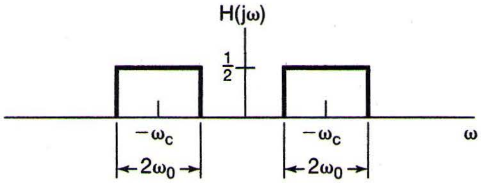


Figure 4.30 Equivalent bandpass filter for $\Re e\{f(t)\}$ in Figure 4.29.

4.6 TABLES OF FOURIER PROPERTIES AND OF BASIC FOURIER TRANSFORM PAIRS

In the preceding sections and in the problems at the end of the chapter, we have considered some of the important properties of the Fourier transform. These are summarized in Table 4.1, in which we have also indicated the section of this chapter in which each property has been discussed.

In Table 4.2, we have assembled a list of many of the basic and important Fourier transform pairs. We will encounter many of these repeatedly as we apply the tools of

TABLE 4.1 PROPERTIES OF THE FOURIER TRANSFORM

Section	Property	Aperiodic signal	Fourier transform
		$x(t)$ $y(t)$	$X(j\omega)$ $Y(j\omega)$
4.3.1	Linearity	$ax(t) + by(t)$	$aX(j\omega) + bY(j\omega)$
4.3.2	Time Shifting	$x(t - t_0)$	$e^{-j\omega t_0}X(j\omega)$
4.3.6	Frequency Shifting	$e^{j\omega_0 t}x(t)$	$X(j(\omega - \omega_0))$
4.3.3	Conjugation	$x^*(t)$	$X^*(-j\omega)$
4.3.5	Time Reversal	$x(-t)$	$X(-j\omega)$
4.3.5	Time and Frequency Scaling	$x(at)$	$\frac{1}{\|a\|}X\left(\frac{j\omega}{a}\right)$
4.4	Convolution	$x(t) * y(t)$	$X(j\omega)Y(j\omega)$
4.5	Multiplication	$x(t)y(t)$	$\frac{1}{2\pi}\int_{-\infty}^{+\infty} X(j\theta)Y(j(\omega-\theta))d\theta$
4.3.4	Differentiation in Time	$\frac{d}{dt}x(t)$	$j\omega X(j\omega)$
4.3.4	Integration	$\int_{-\infty}^{t} x(t)dt$	$\frac{1}{j\omega}X(j\omega) + \pi X(0)\delta(\omega)$
4.3.6	Differentiation in Frequency	$tx(t)$	$j\frac{d}{d\omega}X(j\omega)$
4.3.3	Conjugate Symmetry for Real Signals	$x(t)$ real	$X(j\omega) = X^*(-j\omega)$ $\mathcal{R}e\{X(j\omega)\} = \mathcal{R}e\{X(-j\omega)\}$ $\mathcal{I}m\{X(j\omega)\} = -\mathcal{I}m\{X(-j\omega)\}$ $\|X(j\omega)\| = \|X(-j\omega)\|$ $\sphericalangle X(j\omega) = -\sphericalangle X(-j\omega)$
4.3.3	Symmetry for Real and Even Signals	$x(t)$ real and even	$X(j\omega)$ real and even
4.3.3	Symmetry for Real and Odd Signals	$x(t)$ real and odd	$X(j\omega)$ purely imaginary and odd
4.3.3	Even-Odd Decomposition for Real Signals	$x_e(t) = \mathcal{E}v\{x(t)\}$ [$x(t)$ real] $x_o(t) = \mathcal{O}d\{x(t)\}$ [$x(t)$ real]	$\mathcal{R}e\{X(j\omega)\}$ $j\mathcal{I}m\{X(j\omega)\}$
4.3.7	Parseval's Relation for Aperiodic Signals $\int_{-\infty}^{+\infty} \|x(t)\|^2 dt = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \|X(j\omega)\|^2 d\omega$		

TABLE 4.2 BASIC FOURIER TRANSFORM PAIRS

Signal	Fourier transform	Fourier series coefficients (if periodic)
$\sum_{k=-\infty}^{+\infty} a_k e^{jk\omega_0 t}$	$2\pi \sum_{k=-\infty}^{+\infty} a_k \delta(\omega - k\omega_0)$	a_k
$e^{j\omega_0 t}$	$2\pi\delta(\omega - \omega_0)$	$a_1 = 1$ $a_k = 0$, otherwise
$\cos \omega_0 t$	$\pi[\delta(\omega - \omega_0) + \delta(\omega + \omega_0)]$	$a_1 = a_{-1} = \frac{1}{2}$ $a_k = 0$, otherwise
$\sin \omega_0 t$	$\frac{\pi}{j}[\delta(\omega - \omega_0) - \delta(\omega + \omega_0)]$	$a_1 = -a_{-1} = \frac{1}{2j}$ $a_k = 0$, otherwise
$x(t) = 1$	$2\pi\,\delta(\omega)$	$a_0 = 1, \quad a_k = 0, \ k \neq 0$ (this is the Fourier series representation for any choice of $T > 0$)
Periodic square wave $x(t) = \begin{cases} 1, & \lvert t\rvert < T_1 \\ 0, & T_1 < \lvert t\rvert \leq \frac{T}{2} \end{cases}$ and $x(t + T) = x(t)$	$\sum_{k=-\infty}^{+\infty} \frac{2 \sin k\omega_0 T_1}{k} \delta(\omega - k\omega_0)$	$\frac{\omega_0 T_1}{\pi} \operatorname{sinc}\left(\frac{k\omega_0 T_1}{\pi}\right) = \frac{\sin k\omega_0 T_1}{k\pi}$
$\sum_{n=-\infty}^{+\infty} \delta(t - nT)$	$\frac{2\pi}{T} \sum_{k=-\infty}^{+\infty} \delta\left(\omega - \frac{2\pi k}{T}\right)$	$a_k = \frac{1}{T}$ for all k
$x(t)\begin{cases} 1, & \lvert t\rvert < T_1 \\ 0, & \lvert t\rvert > T_1 \end{cases}$	$\frac{2 \sin \omega T_1}{\omega}$	—
$\frac{\sin Wt}{\pi t}$	$X(j\omega) = \begin{cases} 1, & \lvert\omega\rvert < W \\ 0, & \lvert\omega\rvert > W \end{cases}$	—
$\delta(t)$	1	—
$u(t)$	$\frac{1}{j\omega} + \pi\,\delta(\omega)$	—
$\delta(t - t_0)$	$e^{-j\omega t_0}$	—
$e^{-at}u(t), \mathcal{R}e\{a\} > 0$	$\frac{1}{a + j\omega}$	—
$te^{-at}u(t), \mathcal{R}e\{a\} > 0$	$\frac{1}{(a + j\omega)^2}$	—
$\frac{t^{n-1}}{(n-1)!}e^{-at}u(t)$, $\mathcal{R}e\{a\} > 0$	$\frac{1}{(a + j\omega)^n}$	—

Fourier analysis in our examination of signals and systems. All of the transform pairs, except for the last one in the table, have been considered in examples in the preceding sections. The last pair is considered in Problem 4.40. In addition, note that several of the signals in Table 4.2 are periodic, and for these we have also listed the corresponding Fourier series coefficients.

4.7 SYSTEMS CHARACTERIZED BY LINEAR CONSTANT-COEFFICIENT DIFFERENTIAL EQUATIONS

As we have discussed on several occasions, a particularly important and useful class of continuous-time LTI systems is those for which the input and output satisfy a linear constant-coefficient differential equation of the form

$$\sum_{k=0}^{N} a_k \frac{d^k y(t)}{dt^k} = \sum_{k=0}^{M} b_k \frac{d^k x(t)}{dt^k}. \tag{4.72}$$

In this section, we consider the question of determining the frequency response of such an LTI system. Throughout the discussion we will always assume that the frequency response of the system exists, i.e., that eq. (3.121) converges.

There are two closely related ways in which to determine the frequency response $H(j\omega)$ for an LTI system described by the differential equation (4.72). The first of these, which relies on the fact that complex exponential signals are eigenfunctions of LTI systems, was used in Section 3.10 in our analysis of several simple, nonideal filters. Specifically, if $x(t) = e^{j\omega t}$, then the output must be $y(t) = H(j\omega)e^{j\omega t}$. Substituting these expressions into the differential equation (4.72) and performing some algebra, we can then solve for $H(j\omega)$. In this section we use an alternative approach to arrive at the same answer, making use of the differentiation property, eq. (4.31), of Fourier transforms.

Consider an LTI system characterized by eq. (4.72). From the convolution property,

$$Y(j\omega) = H(j\omega)X(j\omega),$$

or equivalently,

$$H(j\omega) = \frac{Y(j\omega)}{X(j\omega)}, \tag{4.73}$$

where $X(j\omega)$, $Y(j\omega)$, and $H(j\omega)$ are the Fourier transforms of the input $x(t)$, output $y(t)$, and impulse response $h(t)$, respectively. Next, consider applying the Fourier transform to both sides of eq. (4.72) to obtain

$$\mathfrak{F}\left\{\sum_{k=0}^{N} a_k \frac{d^k y(t)}{dt^k}\right\} = \mathfrak{F}\left\{\sum_{k=0}^{M} b_k \frac{d^k x(t)}{dt^k}\right\}. \tag{4.74}$$

From the linearity property, eq. (4.26), this becomes

$$\sum_{k=0}^{N} a_k \mathfrak{F}\left\{\frac{d^k y(t)}{dt^k}\right\} = \sum_{k=0}^{M} b_k \mathfrak{F}\left\{\frac{d^k x(t)}{dt^k}\right\}, \tag{4.75}$$

and from the differentiation property, eq. (4.31),

$$\sum_{k=0}^{N} a_k(j\omega)^k Y(j\omega) = \sum_{k=0}^{M} b_k(j\omega)^k X(j\omega),$$

or equivalently,

$$Y(j\omega)\left[\sum_{k=0}^{N} a_k(j\omega)^k\right] = X(j\omega)\left[\sum_{k=0}^{M} b_k(j\omega)^k\right].$$

Thus, from eq. (4.73),

$$H(j\omega) = \frac{Y(j\omega)}{X(j\omega)} = \frac{\sum_{k=0}^{M} b_k(j\omega)^k}{\sum_{k=0}^{N} a_k(j\omega)^k}. \tag{4.76}$$

Observe that $H(j\omega)$ is thus a rational function; that is, it is a ratio of polynomials in $(j\omega)$. The coefficients of the numerator polynomial are the same coefficients as those that appear on the right-hand side of eq. (4.72), and the coefficients of the denominator polynomial are the same coefficients as appear on the left side of eq. (4.72). Hence, the frequency response given in eq. (4.76) for the LTI system characterized by eq. (4.72) can be written down directly by inspection.

The differential equation (4.72) is commonly referred to as an Nth-order differential equation, as the equation involves derivatives of the output $y(t)$ up through the Nth derivative. Also, the denominator of $H(j\omega)$ in eq. (4.76) is an Nth-order polynomial in $(j\omega)$.

Example 4.24

Consider a stable LTI system characterized by the differential equation

$$\frac{dy(t)}{dt} + ay(t) = x(t), \tag{4.77}$$

with $a > 0$. From eq. (4.76), the frequency response is

$$H(j\omega) = \frac{1}{j\omega + a}. \tag{4.78}$$

Comparing this with the result of Example 4.1, we see that eq. (4.78) is the Fourier transform of $e^{-at}u(t)$. The impulse response of the system is then recognized as

$$h(t) = e^{-at}u(t).$$

Example 4.25

Consider a stable LTI system that is characterized by the differential equation

$$\frac{d^2y(t)}{dt^2} + 4\frac{dy(t)}{dt} + 3y(t) = \frac{dx(t)}{dt} + 2x(t).$$

From eq. (4.76), the frequency response is

$$H(j\omega) = \frac{(j\omega) + 2}{(j\omega)^2 + 4(j\omega) + 3}. \tag{4.79}$$

To determine the corresponding impulse response, we require the inverse Fourier transform of $H(j\omega)$. This can be found using the technique of partial-fraction expansion employed in Example 4.19 and discussed in detail in the appendix. (In particular, see Example A.1, in which the details of the calculations for the partial-fraction expansion of eq. (4.79) are worked out.) As a first step, we factor the denominator of the right-hand side of eq. (4.79) into a product of lower order terms:

$$H(j\omega) = \frac{j\omega + 2}{(j\omega + 1)(j\omega + 3)}. \tag{4.80}$$

Then, using the method of partial-fraction expansion, we find that

$$H(j\omega) = \frac{\frac{1}{2}}{j\omega + 1} + \frac{\frac{1}{2}}{j\omega + 3}.$$

The inverse transform of each term can be recognized from Example 4.24, with the result that

$$h(t) = \frac{1}{2}e^{-t}u(t) + \frac{1}{2}e^{-3t}u(t).$$

The procedure used in Example 4.25 to obtain the inverse Fourier transform is generally useful in inverting transforms that are ratios of polynomials in $j\omega$. In particular, we can use eq. (4.76) to determine the frequency response of any LTI system described by a linear constant-coefficient differential equation and then can calculate the impulse response by performing a partial-fraction expansion that puts the frequency response into a form in which the inverse transform of each term can be recognized by inspection. In addition, if the Fourier transform $X(j\omega)$ of the input to such a system is also a ratio of polynomials in $j\omega$, then so is $Y(j\omega) = H(j\omega)X(j\omega)$. In this case we can use the same technique to solve the differential equation—that is, to find the response $y(t)$ to the input $x(t)$. This is illustrated in the next example.

Example 4.26

Consider the system of Example 4.25, and suppose that the input is

$$x(t) = e^{-t}u(t).$$

Then, using eq. (4.80), we have

$$\begin{aligned} Y(j\omega) = H(j\omega)X(j\omega) &= \left[\frac{j\omega + 2}{(j\omega + 1)(j\omega + 3)}\right]\left[\frac{1}{j\omega + 1}\right] \\ &= \frac{j\omega + 2}{(j\omega + 1)^2(j\omega + 3)}. \end{aligned} \tag{4.81}$$

As discussed in the appendix, in this case the partial-fraction expansion takes the form

$$Y(j\omega) = \frac{A_{11}}{j\omega+1} + \frac{A_{12}}{(j\omega+1)^2} + \frac{A_{21}}{j\omega+3}, \tag{4.82}$$

where A_{11}, A_{12}, and A_{21} are constants to be determined. In Example A.2 in the appendix, the technique of partial-fraction expansion is used to determine these constants. The values obtained are

$$A_{11} = \frac{1}{4}, \quad A_{12} = \frac{1}{2}, \quad A_{21} = -\frac{1}{4},$$

so that

$$Y(j\omega) = \frac{\frac{1}{4}}{j\omega+1} + \frac{\frac{1}{2}}{(j\omega+1)^2} - \frac{\frac{1}{4}}{j\omega+3}. \tag{4.83}$$

Again, the inverse Fourier transform for each term in eq. (4.83) can be obtained by inspection. The first and third terms are of the same type that we have encountered in the preceding two examples, while the inverse transform of the second term can be obtained from Table 4.2 or, as was done in Example 4.19, by applying the dual of the differentiation property, as given in eq. (4.40), to $1/(j\omega+1)$. The inverse transform of eq. (4.83) is then found to be

$$y(t) = \left[\frac{1}{4}e^{-t} + \frac{1}{2}te^{-t} - \frac{1}{4}e^{-3t}\right]u(t).$$

From the preceding examples, we see how the techniques of Fourier analysis allow us to reduce problems concerning LTI systems characterized by differential equations to straightforward algebraic problems. This important fact is illustrated further in a number of the problems at the end of the chapter. In addition (see Chapter 6), the algebraic structure of the rational transforms encountered in dealing with LTI systems described by differential equations greatly facilitate the analysis of their frequency-domain properties and the development of insights into both the time-domain and frequency-domain characteristics of this important class of systems.

4.8 SUMMARY

In this chapter, we have developed the Fourier transform representation for continous-time signals and have examined many of the properties that make this transform so useful. In particular, by viewing an aperiodic signal as the limit of a periodic signal as the period becomes arbitrarily large, we derived the Fourier transform representation for aperiodic signals from the Fourier series representation for periodic signals developed in Chapter 3. In addition, periodic signals themselves can be represented using Fourier transforms consisting of trains of impulses located at the harmonic frequencies of the periodic signal and with areas proportional to the corresponding Fourier series coefficients.

The Fourier transform possesses a wide variety of important properties that describe how different characteristics of signals are reflected in their transforms, and in this chapter we have derived and examined many of these properties. Among them are two that have particular significance for our study of signals and systems. The first is the convolution property, which is a direct consequence of the eigenfunction property of complex exponential signals and which leads to the description of an LTI system in terms of its frequency response. This description plays a fundamental role in the frequency-domain approach to the analysis of LTI systems, which we will continue to explore in subsequent chapters. The second property of the Fourier transform that has extremely important implications is the multiplication property, which provides the basis for the frequency-domain analysis of sampling and modulation systems. We examine these systems further in Chapters 7 and 8.

We have also seen that the tools of Fourier analysis are particularly well suited to the examination of LTI systems characterized by linear constant-coefficient differential equations. Specifically, we have found that the frequency response for such a system can be determined by inspection and that the technique of partial-fraction expansion can then be used to facilitate the calculation of the impulse response of the system. In subsequent chapters, we will find that the convenient algebraic structure of the frequency responses of these systems allows us to gain considerable insight into their characteristics in both the time and frequency domains.

Chapter 4 Problems

The first section of problems belongs to the basic category and the answers are provided in the back of the book. The remaining three sections contain problems belonging to the basic, advanced, and extension categories, respectively.

BASIC PROBLEMS WITH ANSWERS

4.1. Use the Fourier transform analysis equation (4.9) to calculate the Fourier transforms of:

(a) $e^{-2(t-1)}u(t-1)$ **(b)** $e^{-2|t-1|}$

Sketch and label the magnitude of each Fourier transform.

4.2. Use the Fourier transform analysis equation (4.9) to calculate the Fourier transforms of:

(a) $\delta(t+1)+\delta(t-1)$ **(b)** $\frac{d}{dt}\{u(-2-t)+u(t-2)\}$

Sketch and label the magnitude of each Fourier transform.

4.3. Determine the Fourier transform of each of the following periodic signals:

(a) $\sin(2\pi t+\frac{\pi}{4})$ **(b)** $1+\cos(6\pi t+\frac{\pi}{8})$

4.4. Use the Fourier transform synthesis equation (4.8) to determine the inverse Fourier transforms of:

(a) $X_1(j\omega) = 2\pi\,\delta(\omega)+\pi\,\delta(\omega-4\pi)+\pi\,\delta(\omega+4\pi)$

(b) $X_2(j\omega) = \begin{cases} 2, & 0 \le \omega \le 2 \\ -2, & -2 \le \omega < 0 \\ 0, & |\omega| > 2 \end{cases}$

4.5. Use the Fourier transform synthesis equation (4.8) to determine the inverse Fourier transform of $X(j\omega) = |X(j\omega)|e^{j\sphericalangle X(j\omega)}$, where

$$|X(j\omega)| = 2\{u(\omega + 3) - u(\omega - 3)\},$$
$$\sphericalangle X(j\omega) = -\frac{3}{2}\omega + \pi.$$

Use your answer to determine the values of t for which $x(t) = 0$.

4.6. Given that $x(t)$ has the Fourier transform $X(j\omega)$, express the Fourier transforms of the signals listed below in terms of $X(j\omega)$. You may find useful the Fourier transform properties listed in Table 4.1.

(a) $x_1(t) = x(1 - t) + x(-1 - t)$
(b) $x_2(t) = x(3t - 6)$
(c) $x_3(t) = \frac{d^2}{dt^2}x(t - 1)$

4.7. For each of the following Fourier transforms, use Fourier transform properties (Table 4.1) to determine whether the corresponding time-domain signal is (i) real, imaginary, or neither and (ii) even, odd, or neither. Do this without evaluating the inverse of any of the given transforms.

(a) $X_1(j\omega) = u(\omega) - u(\omega - 2)$
(b) $X_2(j\omega) = \cos(2\omega)\sin(\frac{\omega}{2})$
(c) $X_3(j\omega) = A(\omega)e^{jB(\omega)}$, where $A(\omega) = (\sin 2\omega)/\omega$ and $B(\omega) = 2\omega + \frac{\pi}{2}$
(d) $X(j\omega) = \sum_{k=-\infty}^{\infty}(\frac{1}{2})^{|k|}\,\delta(\omega - \frac{k\pi}{4})$

4.8. Consider the signal

$$x(t) = \begin{cases} 0, & t < -\frac{1}{2} \\ t + \frac{1}{2}, & -\frac{1}{2} \le t \le \frac{1}{2} \\ 1, & t > \frac{1}{2} \end{cases}.$$

(a) Use the differentiation and integration properties in Table 4.1 and the Fourier transform pair for the rectangular pulse in Table 4.2 to find a closed-form expression for $X(j\omega)$.
(b) What is the Fourier transform of $g(t) = x(t) - \frac{1}{2}$?

4.9. Consider the signal

$$x(t) = \begin{cases} 0, & |t| > 1 \\ (t + 1)/2, & -1 \le t \le 1 \end{cases}.$$

(a) With the help of Tables 4.1 and 4.2, determine the closed-form expression for $X(j\omega)$.
(b) Take the real part of your answer to part (a), and verify that it is the Fourier transform of the even part of $x(t)$.
(c) What is the Fourier transform of the odd part of $x(t)$?

4.10. **(a)** Use Tables 4.1 and 4.2 to help determine the Fourier transform of the following signal:

$$x(t) = t\left(\frac{\sin t}{\pi t}\right)^2$$

(b) Use Parseval's relation and the result of the previous part to determine the numerical value of

$$A = \int_{-\infty}^{+\infty} t^2 \left(\frac{\sin t}{\pi t}\right)^4 dt$$

4.11. Given the relationships

$$y(t) = x(t) * h(t)$$

and

$$g(t) = x(3t) * h(3t),$$

and given that $x(t)$ has Fourier transform $X(j\omega)$ and $h(t)$ has Fourier transform $H(j\omega)$, use Fourier transform properties to show that $g(t)$ has the form

$$g(t) = Ay(Bt).$$

Determine the values of A and B.

4.12. Consider the Fourier transform pair

$$e^{-|t|} \overset{\mathcal{F}}{\longleftrightarrow} \frac{2}{1+\omega^2}.$$

(a) Use the appropriate Fourier transform properties to find the Fourier transform of $te^{-|t|}$.

(b) Use the result from part (a), along with the duality property, to determine the Fourier transform of

$$\frac{4t}{(1+t^2)^2}.$$

Hint: See Example 4.13.

4.13. Let $x(t)$ be a signal whose Fourier transform is

$$X(j\omega) = \delta(\omega) + \delta(\omega - \pi) + \delta(\omega - 5),$$

and let

$$h(t) = u(t) - u(t-2).$$

(a) Is $x(t)$ periodic?

(b) Is $x(t) * h(t)$ periodic?

(c) Can the convolution of two aperiodic signals be periodic?

4.14. Consider a signal $x(t)$ with Fourier transform $X(j\omega)$. Suppose we are given the following facts:

1. $x(t)$ is real and nonnegative.
2. $\mathfrak{F}^{-1}\{(1 + j\omega)X(j\omega)\} = Ae^{-2t}u(t)$, where A is independent of t.
3. $\int_{-\infty}^{\infty} |X(j\omega)|^2 \, d\omega = 2\pi$.

Determine a closed-form expression for $x(t)$.

4.15. Let $x(t)$ be a signal with Fourier transform $X(j\omega)$. Suppose we are given the following facts:

1. $x(t)$ is real.
2. $x(t) = 0$ for $t \leq 0$.
3. $\frac{1}{2\pi}\int_{-\infty}^{\infty} \mathcal{R}e\{X(j\omega)\}e^{j\omega t} \, d\omega = |t|e^{-|t|}$.

Determine a closed-form expression for $x(t)$.

4.16. Consider the signal

$$x(t) = \sum_{k=-\infty}^{\infty} \frac{\sin(k\frac{\pi}{4})}{(k\frac{\pi}{4})} \delta(t - k\frac{\pi}{4}).$$

(a) Determine $g(t)$ such that

$$x(t) = \left(\frac{\sin t}{\pi t}\right) g(t).$$

(b) Use the multiplication property of the Fourier transform to argue that $X(j\omega)$ is periodic. Specify $X(j\omega)$ over one period.

4.17. Determine whether each of the following statements is true or false. Justify your answers.

(a) An odd and imaginary signal always has an odd and imaginary Fourier transform.

(b) The convolution of an odd Fourier transform with an even Fourier transform is always odd.

4.18. Find the impulse response of a system with the frequency response

$$H(j\omega) = \frac{(\sin^2(3\omega)) \cos \omega}{\omega^2}.$$

4.19. Consider a causal LTI system with frequency response

$$H(j\omega) = \frac{1}{j\omega + 3}.$$

For a particular input $x(t)$ this system is observed to produce the output

$$y(t) = e^{-3t}u(t) - e^{-4t}u(t).$$

Determine $x(t)$.

4.20. Find the impulse response of the causal LTI system represented by the *RLC* circuit considered in Problem 3.20. Do this by taking the inverse Fourier transform of the circuit's frequency response. You may use Tables 4.1 and 4.2 to help evaluate the inverse Fourier transform.

BASIC PROBLEMS

4.21. Compute the Fourier transform of each of the following signals:

(a) $[e^{-\alpha t}\cos \omega_0 t]u(t),\ \alpha > 0$

(b) $e^{-3|t|}\sin 2t$

(c) $x(t) = \begin{cases} 1 + \cos \pi t, & |t| \le 1 \\ 0, & |t| > 1 \end{cases}$

(d) $\sum_{k=0}^{\infty} \alpha^k \delta(t - kT),\ |\alpha| < 1$

(e) $[te^{-2t}\sin 4t]u(t)$

(f) $\left[\frac{\sin \pi t}{\pi t}\right]\left[\frac{\sin 2\pi(t-1)}{\pi(t-1)}\right]$

(g) $x(t)$ as shown in Figure P4.21(a)

(h) $x(t)$ as shown in Figure P4.21(b)

(i) $x(t) = \begin{cases} 1 - t^2, & 0 < t < 1 \\ 0, & \text{otherwise} \end{cases}$

(j) $\sum_{n=-\infty}^{+\infty} e^{-|t-2n|}$

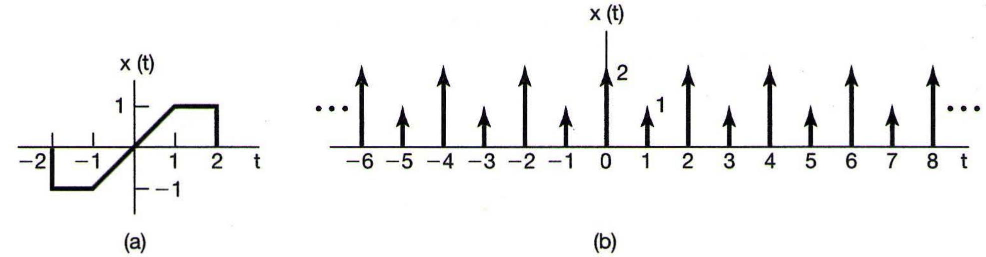


Figure P4.21

4.22. Determine the continuous-time signal corresponding to each of the following transforms.

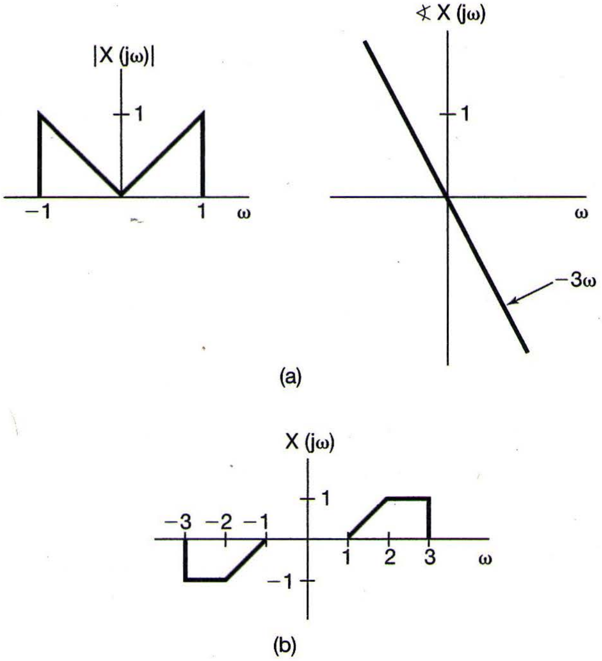


Figure P4.22

(a) $X(j\omega) = \frac{2\sin[3(\omega-2\pi)]}{(\omega-2\pi)}$
(b) $X(j\omega) = \cos(4\omega + \pi/3)$
(c) $X(j\omega)$ as given by the magnitude and phase plots of Figure P4.22(a)
(d) $X(j\omega) = 2[\delta(\omega - 1) - \delta(\omega + 1)] + 3[\delta(\omega - 2\pi) + \delta(\omega + 2\pi)]$
(e) $X(j\omega)$ as in Figure P4.22(b)

4.23. Consider the signal

$$x_0(t) = \begin{cases} e^{-t}, & 0 \le t \le 1 \\ 0, & \text{elsewhere} \end{cases}.$$

Determine the Fourier transform of each of the signals shown in Figure P4.23. You should be able to do this by explicitly evaluating *only* the transform of $x_0(t)$ and then using properties of the Fourier transform.

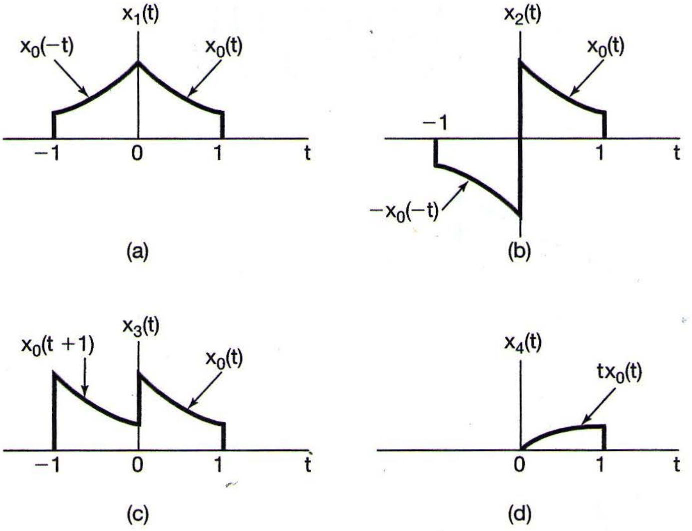


Figure P4.23

4.24. **(a)** Determine which, if any, of the real signals depicted in Figure P4.24 have Fourier transforms that satisfy each of the following conditions:
(1) $\mathcal{R}e\{X(j\omega)\} = 0$
(2) $\mathcal{I}m\{X(j\omega)\} = 0$
(3) There exists a real α such that $e^{j\alpha\omega}X(j\omega)$ is real
(4) $\int_{-\infty}^{\infty} X(j\omega)\, d\omega = 0$
(5) $\int_{-\infty}^{\infty} \omega X(j\omega)\, d\omega = 0$
(6) $X(j\omega)$ is periodic

(b) Construct a signal that has properties (1), (4), and (5) and does *not* have the others.

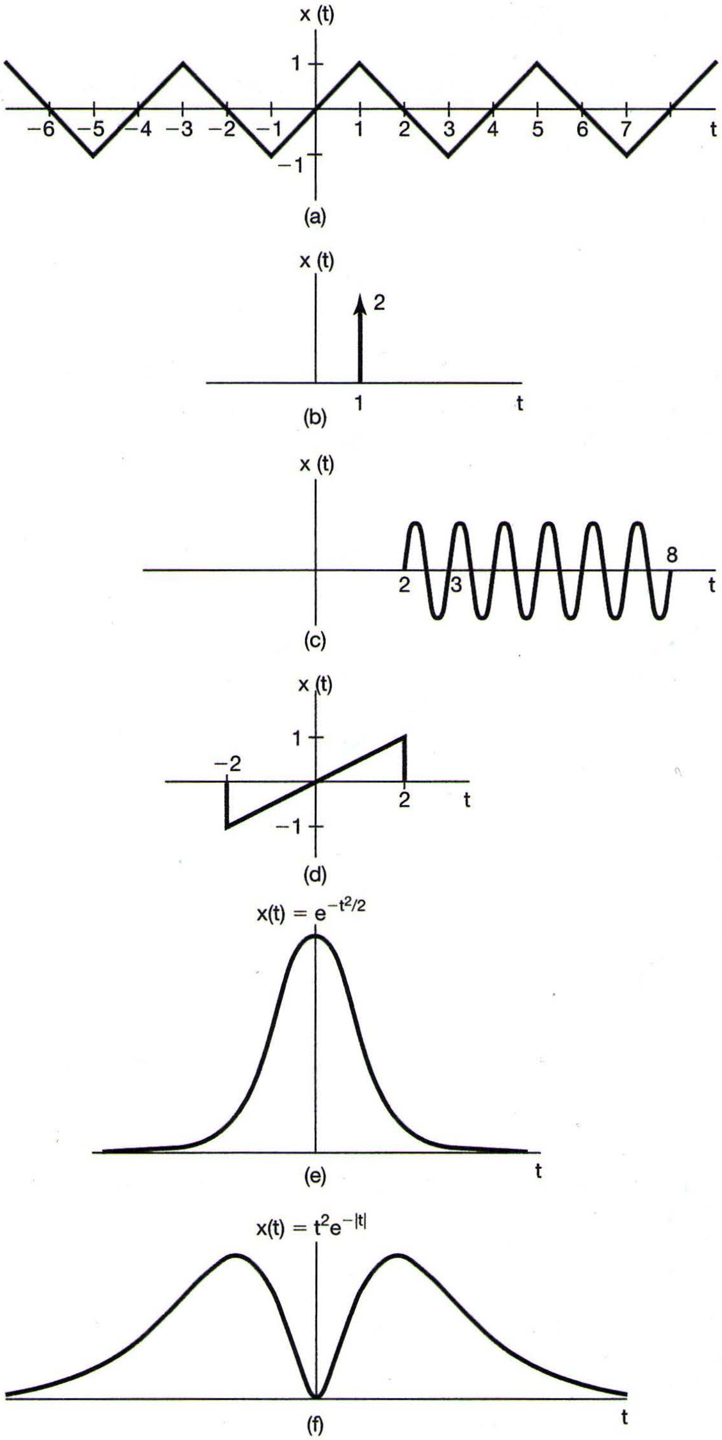

x (t)
1
−6
−5
−4
−3
−2
−1
1
2
3
4
5
6
7
t
−1
(a)
x (t)
2
1
t
(b)
x (t)
8
2
3
t
(c)
x (t)
1
−2
2
t
−1
(d)
x(t) = $e^{-t^2/2}$
t
(e)
x(t) = $t^2e^{-|t|}$
t
(f)

Figure P4.24

4.25. Let $X(j\omega)$ denote the Fourier transform of the signal $x(t)$ depicted in Figure P4.25.

(a) $X(j\omega)$ can be expressed as $A(j\omega)e^{j\Theta(j\omega)}$, where $A(j\omega)$ and $\Theta(j\omega)$ are both real-values. Find $\Theta(j\omega)$.

(b) Find $X(j0)$.

(c) Find $\int_{-\infty}^{\infty} X(j\omega)\,d\omega$.

(d) Evaluate $\int_{-\infty}^{\infty} X(j\omega)\frac{2\sin\omega}{\omega}e^{j2\omega}\,d\omega$.

(e) Evaluate $\int_{-\infty}^{\infty} |X(j\omega)|^2\,d\omega$.

(f) Sketch the inverse Fourier transform of $\mathcal{R}e\{X(j\omega)\}$.

Note: You should perform all these calculations without explicitly evaluating $X(j\omega)$.

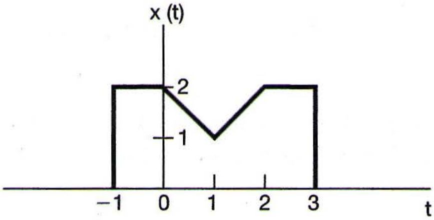


Figure P4.25

4.26. **(a)** Compute the convolution of each of the following pairs of signals $x(t)$ and $h(t)$ by calculating $X(j\omega)$ and $H(j\omega)$, using the convolution property, and inverse transforming.

(i) $x(t) = te^{-2t}u(t),\ h(t) = e^{-4t}u(t)$

(ii) $x(t) = te^{-2t}u(t),\ h(t) = te^{-4t}u(t)$

(iii) $x(t) = e^{-t}u(t),\ h(t) = e^{t}u(-t)$

(b) Suppose that $x(t) = e^{-(t-2)}u(t-2)$ and $h(t)$ is as depicted in Figure P4.26. Verify the convolution property for this pair of signals by showing that the Fourier transform of $y(t) = x(t) * h(t)$ equals $H(j\omega)X(j\omega)$.

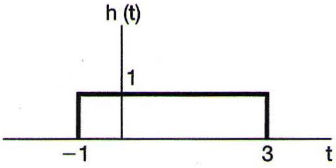


Figure P4.26

4.27. Consider the signals

$$x(t) = u(t-1) - 2u(t-2) + u(t-3)$$

and

$$\tilde{x}(t) = \sum_{k=-\infty}^{\infty} x(t-kT),$$

where $T > 0$. Let a_k denote the Fourier series coefficients of $\tilde{x}(t)$, and let $X(j\omega)$ denote the Fourier transform of $x(t)$.

(a) Determine a closed-form expression for $X(j\omega)$.

(b) Determine an expression for the Fourier coefficients a_k and verify that $a_k = \frac{1}{T}X\left(j\frac{2\pi k}{T}\right)$.

4.28. **(a)** Let $x(t)$ have the Fourier transform $X(j\omega)$, and let $p(t)$ be periodic with fundamental frequency ω_0 and Fourier series representation

$$p(t) = \sum_{n=-\infty}^{+\infty} a_n e^{jn\omega_0 t}.$$

Determine an expression for the Fourier transform of

$$y(t) = x(t)p(t). \qquad \text{(P4.28–1)}$$

(b) Suppose that $X(j\omega)$ is as depicted in Figure P4.28(a). Sketch the spectrum of $y(t)$ in eq. (P4.28–1) for each of the following choices of $p(t)$:

(i) $p(t) = \cos(t/2)$
(ii) $p(t) = \cos t$
(iii) $p(t) = \cos 2t$
(iv) $p(t) = (\sin t)(\sin 2t)$
(v) $p(t) = \cos 2t - \cos t$
(vi) $p(t) = \sum_{n=-\infty}^{+\infty} \delta(t - \pi n)$
(vii) $p(t) = \sum_{n=-\infty}^{+\infty} \delta(t - 2\pi n)$
(viii) $p(t) = \sum_{n=-\infty}^{+\infty} \delta(t - 4\pi n)$
(ix) $p(t) = \sum_{n=-\infty}^{+\infty} \delta(t - 2\pi n) - \frac{1}{2}\sum_{n=-\infty}^{+\infty} \delta(t - \pi n)$
(x) $p(t) =$ the periodic square wave shown in Figure P4.28(b).

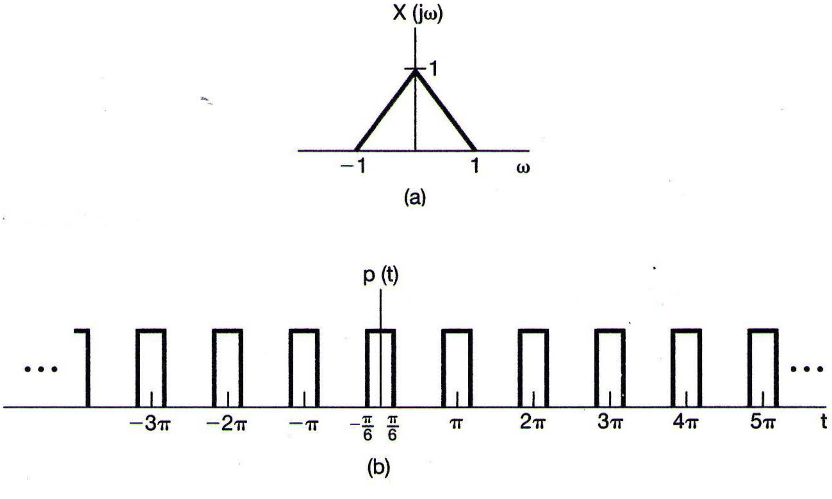


Figure P4.28

4.29. A real-valued continuous-time function $x(t)$ has a Fourier transform $X(j\omega)$ whose magnitude and phase are as illustrated in Figure P4.29(a).

The functions $x_a(t)$, $x_b(t)$, $x_c(t)$, and $x_d(t)$ have Fourier transforms whose magnitudes are identical to $X(j\omega)$, but whose phase functions differ, as shown in Figures P4.29(b)–(e). The phase functions $\sphericalangle X_a(j\omega)$ and $\sphericalangle X_b(j\omega)$ are formed by adding a linear phase to $\sphericalangle X(j\omega)$. The function $\sphericalangle X_c(j\omega)$ is formed by reflecting $\sphericalangle X(j\omega)$ about $\omega = 0$, and $\sphericalangle X_d(j\omega)$ is obtained by a combination of a reflection and an addition of a linear phase. Using the properties of Fourier transforms, determine the expressions for $x_a(t)$, $x_b(t)$, $x_c(t)$, and $x_d(t)$ in terms of $x(t)$.

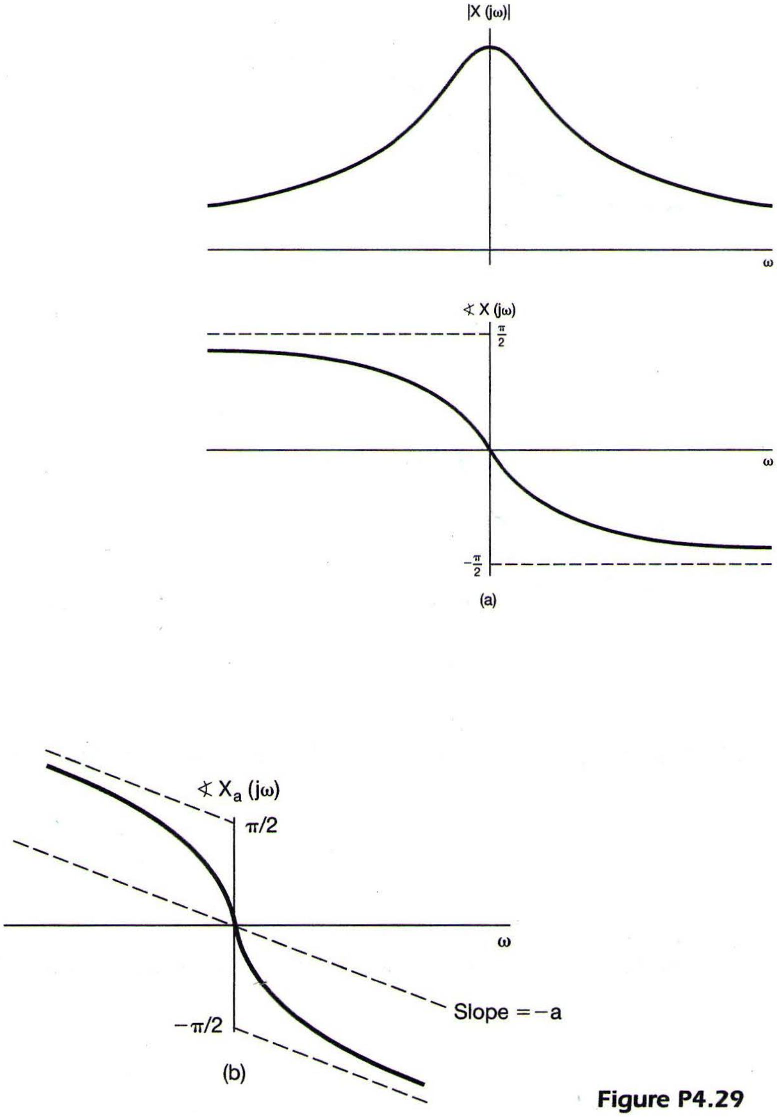


Figure P4.29

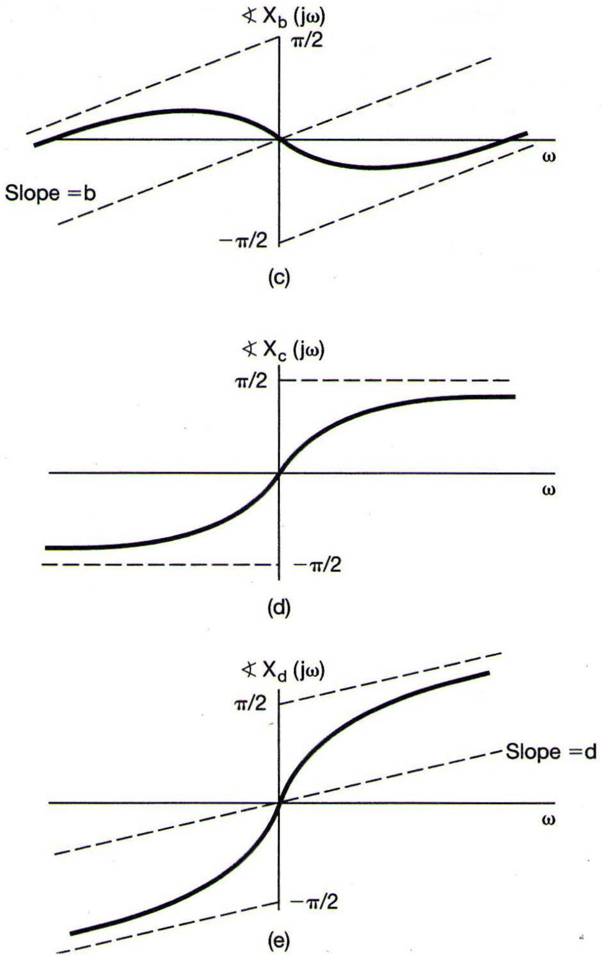


Figure P4.29 *Continued*

4.30. Suppose $g(t) = x(t)\cos t$ and the Fourier transform of the $g(t)$ is

$$G(j\omega) = \begin{cases} 1, & |\omega| \leq 2 \\ 0, & \text{otherwise} \end{cases}.$$

(a) Determine $x(t)$.

(b) Specify the Fourier transform $X_1(j\omega)$ of a signal $x_1(t)$ such that

$$g(t) = x_1(t)\cos\left(\frac{2}{3}t\right).$$

4.31. **(a)** Show that the three LTI systems with impulse responses

$$h_1(t) = u(t),$$

$$h_2(t) = -2\delta(t) + 5e^{-2t}u(t),$$

and

$$h_3(t) = 2te^{-t}u(t)$$

all have the same response to $x(t) = \cos t$.

(b) Find the impulse response of another LTI system with the same response to $\cos t$.

This problem illustrates the fact that the response to $\cos t$ cannot be used to specify an LTI system uniquely.

4.32. Consider an LTI system S with impulse response

$$h(t) = \frac{\sin(4(t-1))}{\pi(t-1)}.$$

Determine the output of S for each of the following inputs:

(a) $x_1(t) = \cos(6t + \frac{\pi}{2})$

(b) $x_2(t) = \sum_{k=0}^{\infty}(\frac{1}{2})^k \sin(3kt)$

(c) $x_3(t) = \frac{\sin(4(t+1))}{\pi(t+1)}$

(d) $x_4(t) = (\frac{\sin 2t}{\pi t})^2$

4.33. The input and the output of a stable and causal LTI system are related by the differential equation

$$\frac{d^2y(t)}{dt^2} + 6\frac{dy(t)}{dt} + 8y(t) = 2x(t)$$

(a) Find the impulse response of this system.

(b) What is the response of this system if $x(t) = te^{-2t}u(t)$?

(c) Repeat part (a) for the stable and causal LTI system described by the equation

$$\frac{d^2y(t)}{dt^2} + \sqrt{2}\frac{dy(t)}{dt} + y(t) = 2\frac{d^2x(t)}{dt^2} - 2x(t)$$

4.34. A causal and stable LTI system S has the frequency response

$$H(j\omega) = \frac{j\omega + 4}{6 - \omega^2 + 5j\omega}.$$

(a) Determine a differential equation relating the input $x(t)$ and output $y(t)$ of S.
(b) Determine the impulse response $h(t)$ of S.
(c) What is the output of S when the input is

$$x(t) = e^{-4t}u(t) - te^{-4t}u(t)?$$

4.35. In this problem, we provide examples of the effects of nonlinear changes in phase.

(a) Consider the continuous-time LTI system with frequency response

$$H(j\omega) = \frac{a - j\omega}{a + j\omega},$$

where $a > 0$. What is the magnitude of $H(j\omega)$? What is $\sphericalangle H(j\omega)$? What is the impulse response of this system?

(b) Determine the output of the system of part (a) with $a = 1$ when the input is

$$\cos(t/\sqrt{3}) + \cos t + \cos \sqrt{3}t.$$

Roughly sketch both the input and the output.

4.36. Consider an LTI system whose response to the input

$$x(t) = [e^{-t} + e^{-3t}]u(t)$$

is

$$y(t) = [2e^{-t} - 2e^{-4t}]u(t).$$

(a) Find the frequency response of this system.
(b) Determine the system's impulse response.
(c) Find the differential equation relating the input and the output of this system.

ADVANCED PROBLEMS

4.37. Consider the signal $x(t)$ in Figure P4.37.

(a) Find the Fourier transform $X(j\omega)$ of $x(t)$.
(b) Sketch the signal

$$\tilde{x}(t) = x(t) * \sum_{k=-\infty}^{\infty} \delta(t - 4k).$$

(c) Find another signal $g(t)$ such that $g(t)$ is not the same as $x(t)$ and

$$\tilde{x}(t) = g(t) * \sum_{k=-\infty}^{\infty} \delta(t - 4k).$$

(d) Argue that, although $G(j\omega)$ is different from $X(j\omega)$, $G(j\frac{\pi k}{2}) = X(j\frac{\pi k}{2})$ for all integers k. You should not explicitly evaluate $G(j\omega)$ to answer this question.

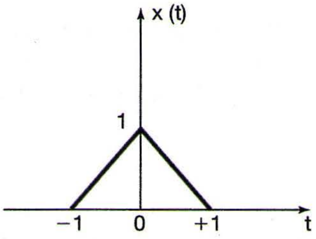


Figure P4.37

4.38. Let $x(t)$ be any signal with Fourier transform $X(j\omega)$. The frequency-shift property of the Fourier transform may be stated as

$$e^{j\omega_0 t}x(t) \overset{\mathcal{F}}{\longleftrightarrow} X(j(\omega - \omega_0)).$$

(a) Prove the frequency-shift property by applying the frequency shift to the analysis equation

$$X(j\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}\,dt.$$

(b) Prove the frequency-shift property by utilizing the Fourier transform of $e^{j\omega_0 t}$ in conjunction with the multiplication property of the Fourier transform.

4.39. Suppose that a signal $x(t)$ has Fourier transform $X(j\omega)$. Now consider another signal $g(t)$ whose shape is the same as the shape of $X(j\omega)$; that is,

$$g(t) = X(jt).$$

(a) Show that the Fourier transform $G(j\omega)$ of $g(t)$ has the same shape as $2\pi x(-t)$; that is, show that

$$G(j\omega) = 2\pi x(-\omega).$$

(b) Using the fact that

$$\mathcal{F}\{\delta(t + B)\} = e^{jB\omega}$$

in conjunction with the result from part (a), show that

$$\mathcal{F}\{e^{jBt}\} = 2\pi\,\delta(\omega - B).$$

4.40. Use properties of the Fourier transform to show by induction that the Fourier transform of

$$x(t) = \frac{t^{n-1}}{(n-1)!}e^{-at}u(t),\ a > 0,$$

is

$$\frac{1}{(a + j\omega)^n}.$$

4.41. In this problem, we derive the multiplication property of the continuous-time Fourier transform. Let $x(t)$ and $y(t)$ be two continuous-time signals with Fourier transforms $X(j\omega)$ and $Y(j\omega)$, respectively. Also, let $g(t)$ denote the inverse Fourier transform of $\frac{1}{2\pi}\{X(j\omega) * Y(j\omega)\}$.

(a) Show that

$$g(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} X(j\theta)\left[\frac{1}{2\pi}\int_{-\infty}^{+\infty} Y(j(\omega - \theta))e^{j\omega t}\,d\omega\right]d\theta.$$

(b) Show that

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} Y(j(\omega - \theta))e^{j\omega t}\,d\omega = e^{j\theta t}y(t).$$

(c) Combine the results of parts (a) and (b) to conclude that

$$g(t) = x(t)y(t).$$

4.42. Let

$$g_1(t) = \{[\cos(\omega_0 t)]x(t)\} * h(t) \qquad \text{and} \qquad g_2(t) = \{[\sin(\omega_0 t)]x(t)\} * h(t),$$

where

$$x(t) = \sum_{k=-\infty}^{\infty} a_k e^{jk100t}$$

is a real-valued periodic signal and $h(t)$ is the impulse response of a stable LTI system.

(a) Specify a value for ω_0 and any necessary constraints on $H(j\omega)$ to ensure that

$$g_1(t) = \mathcal{R}e\{a_5\} \qquad \text{and} \qquad g_2(t) = \mathcal{I}m\{a_5\}.$$

(b) Give an example of $h(t)$ such that $H(j\omega)$ satisfies the constraints you specified in part (a).

4.43. Let

$$g(t) = x(t)\cos^2 t * \frac{\sin t}{\pi t}.$$

Assuming that $x(t)$ is real and $X(j\omega) = 0$ for $|\omega| \geq 1$, show that there exists an LTI system S such that

$$x(t) \xrightarrow{S} g(t).$$

4.44. The output $y(t)$ of a causal LTI system is related to the input $x(t)$ by the equation

$$\frac{dy(t)}{dt} + 10y(t) = \int_{-\infty}^{+\infty} x(\tau)z(t-\tau)\,d\tau - x(t),$$

where $z(t) = e^{-t}u(t) + 3\,\delta(t)$.

(a) Find the frequency response $H(j\omega) = Y(j\omega)/X(j\omega)$ of this system.

(b) Determine the impulse response of the system.

4.45. In the discussion in Section 4.3.7 of Parseval's relation for continuous-time signals, we saw that

$$\int_{-\infty}^{+\infty} |x(t)|^2\,dt = \frac{1}{2\pi}\int_{-\infty}^{+\infty} |X(j\omega)|^2\,d\omega.$$

This says that the total energy of the signal can be obtained by integrating $|X(j\omega)|^2$ over all frequencies. Now consider a real-valued signal $x(t)$ processed by the ideal bandpass filter $H(j\omega)$ shown in Figure P4.45. Express the energy in the output signal $y(t)$ as an integration over frequency of $|X(j\omega)|^2$. For Δ sufficiently small so that $|X(j\omega)|$ is approximately constant over a frequency interval of width Δ, show that the energy in the output $y(t)$ of the bandpass filter is approximately proportional to $\Delta|X(j\omega_0)|^2$.

On the basis of the foregoing result, $\Delta|X(j\omega_0)|^2$ is proportional to the energy in the signal in a bandwidth Δ around the frequency ω_0. For this reason, $|X(j\omega)|^2$ is often referred to as the *energy-density spectrum* of the signal $x(t)$.

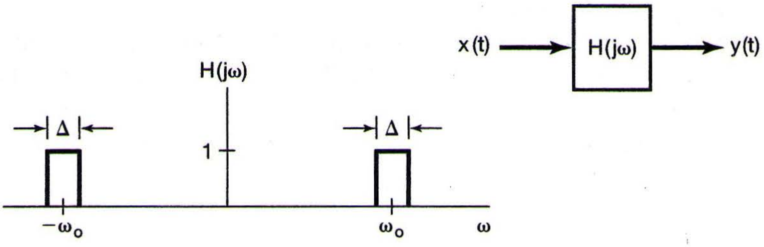


Figure P4.45

4.46. In Section 4.5.1, we discussed the use of amplitude modulation with a complex exponential carrier to implement a bandpass filter. The specific system was shown in Figure 4.26, and if only the real part of $f(t)$ is retained, the equivalent bandpass filter is that shown in Figure 4.30.

In Figure P4.46, we indicate an implementation of a bandpass filter using sinusoidal modulation and lowpass filters. Show that the output $y(t)$ of the system is identical to that which would be obtained by retaining only $\mathcal{R}e\{f(t)\}$ in Figure 4.26.

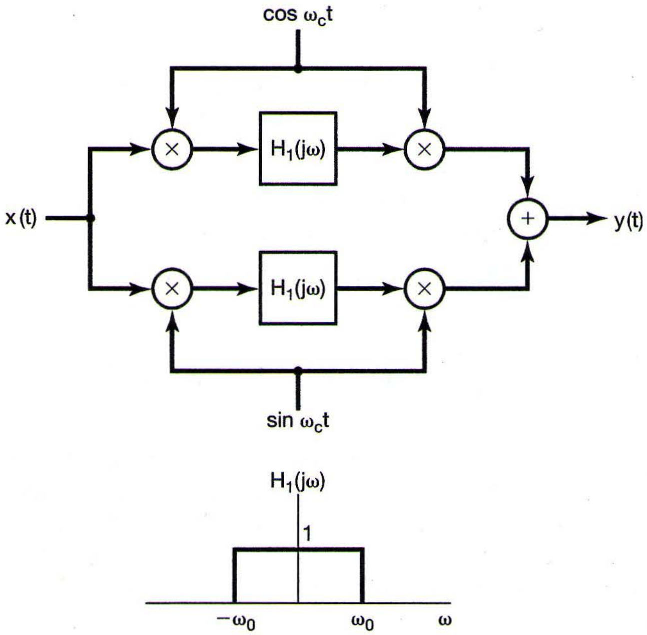


Figure P4.46

4.47. An important property of the frequency response $H(j\omega)$ of a continuous-time LTI system with a real, causal impulse response $h(t)$ is that $H(j\omega)$ is completely specified by its real part, $\mathcal{R}e\{H(j\omega)\}$. The current problem is concerned with deriving and examining some of the implications of this property, which is generally referred to as *real-part sufficiency*.

(a) Prove the property of real-part sufficiency by examining the signal $h_e(t)$, which is the even part of $h(t)$. What is the Fourier transform of $h_e(t)$? Indicate how $h(t)$ can be recovered from $h_e(t)$.

(b) If the real part of the frequency response of a causal system is

$$\mathcal{R}e\{H(j\omega)\} = \cos\omega,$$

what is $h(t)$?

(c) Show that $h(t)$ can be recovered from $h_o(t)$, the odd part of $h(t)$, for every value of t except $t = 0$. Note that if $h(t)$ does not contain any singularities [$\delta(t)$, $u_1(t)$, $u_2(t)$, etc.] at $t = 0$, then the frequency response

$$H(j\omega) = \int_{-\infty}^{+\infty} h(t)e^{-j\omega t}\,dt$$

will not change if $h(t)$ is set to some arbitrary finite value at the single point $t = 0$. Thus, in this case, show that $H(j\omega)$ is also completely specified by its imaginary part.

EXTENSION PROBLEMS

4.48. Let us consider a system with a real and causal impulse response $h(t)$ that does not have any singularities at $t = 0$. In Problem 4.47, we saw that either the real or the imaginary part of $H(j\omega)$ completely determines $H(j\omega)$. In this problem we derive an explicit relationship between $H_R(j\omega)$ and $H_I(j\omega)$, the real and imaginary parts of $H(j\omega)$.

(a) To begin, note that since $h(t)$ is causal,

$$h(t) = h(t)u(t), \tag{P4.48–1}$$

except perhaps at $t = 0$. Now, since $h(t)$ contains no singularities at $t = 0$, the Fourier transforms of both sides of eq. (P4.48–1) must be identical. Use this fact, together with the multiplication property, to show that

$$H(j\omega) = \frac{1}{j\pi}\int_{-\infty}^{+\infty} \frac{H(j\eta)}{\omega - \eta}d\eta. \tag{P4.48–2}$$

Use eq. (P4.48–2) to determine an expression for $H_R(j\omega)$ in terms of $H_I(j\omega)$ and one for $H_I(j\omega)$ in terms of $H_R(j\omega)$.

(b) The operation

$$y(t) = \frac{1}{\pi}\int_{-\infty}^{+\infty} \frac{x(\tau)}{t - \tau}\,d\tau \tag{P4.48–3}$$

is called the *Hilbert transform*. We have just seen that the real and imaginary parts of the transform of a real, causal impulse response $h(t)$ can be determined from one another using the Hilbert transform.

Now consider eq. (P4.48–3), and regard $y(t)$ as the output of an LTI system with input $x(t)$. Show that the frequency response of this system is

$$H(j\omega) = \begin{cases} -j, & \omega > 0 \\ j, & \omega < 0 \end{cases}.$$

(c) What is the Hilbert transform of the signal $x(t) = \cos 3t$?

4.49. Let $H(j\omega)$ be the frequency response of a continuous-time LTI system, and suppose that $H(j\omega)$ is real, even, and positive. Also, assume that

$$\max_{\omega}\{H(j\omega)\} = H(0).$$

(a) Show that:

(i) The impulse response, $h(t)$, is real.

(ii) $\max\{|h(t)|\} = h(0)$.

Hint: If $f(t, \omega)$ is a complex function of two variables, then

$$\left|\int_{-\infty}^{+\infty} f(t,\omega)\,d\omega\right| \le \int_{-\infty}^{+\infty} |f(t,\omega)|\,d\omega.$$

(b) One important concept in system analysis is the *bandwidth* of an LTI system. There are many different mathematical ways in which to define bandwidth, but they are related to the qualitative and intuitive idea that a system with frequency response $G(j\omega)$ essentially "stops" signals of the form $e^{j\omega t}$ for values of ω where $G(j\omega)$ vanishes or is small and "passes" those complex exponentials in the band of frequency where $G(j\omega)$ is not small. The width of this band is the bandwidth. These ideas will be made much clearer in Chapter 6, but for now we will consider a special definition of bandwidth for those systems with frequency responses that have the properties specified previously for $H(j\omega)$. Specifically, one definition of the bandwidth B_w of such a system is the width of the rectangle of height $H(j0)$ that has an area equal to the area under $H(j\omega)$. This is illustrated in Figure P4.49(a). Note that since $H(j0) = \max_\omega H(j\omega)$, the frequencies within the band indicated in the figure are those for which $H(j\omega)$ is largest. The exact choice of the width in the figure is, of course, a bit arbitrary, but we have chosen one definition that allows us to compare different systems and to make precise a very important relationship between time and frequency.

What is the bandwidth of the system with frequency response

$$H(j\omega) = \begin{cases} 1, & |\omega| < W \\ 0, & |\omega| > W \end{cases}?$$

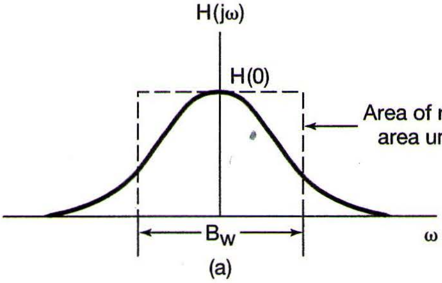


Figure P4.49a

(c) Find an expression for the bandwidth B_w in terms of $H(j\omega)$.

(d) Let $s(t)$ denote the step response of the system set out in part (a). An important measure of the speed of response of a system is the *rise time*, which, like the bandwidth, has a qualitative definition, leading to many possible mathematical definitions, one of which we will use. Intuitively, the rise time of a system is a measure of how fast the step response rises from zero to its final value,

$$s(\infty) = \lim_{t\to\infty} s(t).$$

Thus, the smaller the rise time, the faster is the response of the system. For the system under consideration in this problem, we will define the rise time as

$$t_r = \frac{s(\infty)}{h(0)}.$$

Since

$$s'(t) = h(t),$$

and also because of the property that $h(0) = \max_t h(t)$, t_r is the time it would take to go from zero to $s(\infty)$ while maintaining the maximum rate of change of $s(t)$. This is illustrated in Figure P4.49(b).

Find an expression for t_r in terms of $H(j\omega)$.

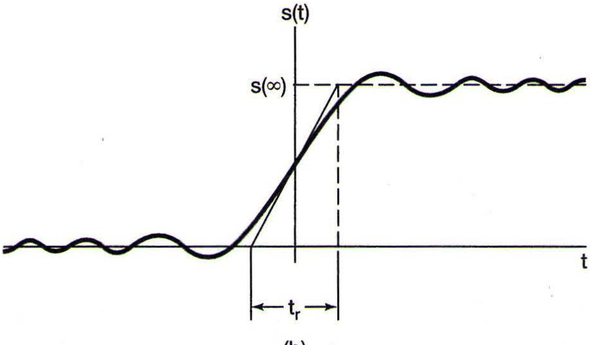


Figure P4.49b

(e) Combine the results of parts (c) and (d) to show that

$$B_w t_r = 2\pi. \tag{P4.49–1}$$

Thus, we *cannot* independently specify both the rise time and the bandwidth of our system. For example, eq. (P4.49–1) implies that, if we want a fast system (t_r small), the system must have a large bandwidth. This is a fundamental trade-off that is of central importance in many problems of system design.

4.50. In Problems 1.45 and 2.67, we defined and examined several of the properties and uses of correlation functions. In the current problem, we examine the properties of such functions in the frequency domain. Let $x(t)$ and $y(t)$ be two real signals. Then the cross-correlation function of $x(t)$ and $y(t)$ is defined as

$$\phi_{xy}(t) = \int_{-\infty}^{+\infty} x(t+\tau)y(\tau)\,d\tau.$$

Similarly, we can define $\phi_{yx}(t)$, $\phi_{xx}(t)$, and $\phi_{yy}(t)$. [The last two of these are called the autocorrelation functions of the signals $x(t)$ and $y(t)$, respectively.] Let $\Phi_{xy}(j\omega)$, $\Phi_{yx}(j\omega)$, $\Phi_{xx}(j\omega)$, and $\Phi_{yy}(j\omega)$ denote the Fourier transforms of $\phi_{xy}(t)$, $\phi_{yx}(t)$, $\phi_{xx}(t)$, and $\phi_{yy}(t)$, respectively.

(a) What is the relationship between $\Phi_{xy}(j\omega)$ and $\Phi_{yx}(j\omega)$?

(b) Find an expression for $\Phi_{xy}(j\omega)$ in terms of $X(j\omega)$ and $Y(j\omega)$.

(c) Show that $\Phi_{xx}(j\omega)$ is real and nonnegative for every ω.

(d) Suppose now that $x(t)$ is the input to an LTI system with a real-valued impulse response and with frequency response $H(j\omega)$ and that $y(t)$ is the output. Find expressions for $\Phi_{xy}(j\omega)$ and $\Phi_{yy}(j\omega)$ in terms of $\Phi_{xx}(j\omega)$ and $H(j\omega)$.

(e) Let $x(t)$ be as is illustrated in Figure P4.50, and let the LTI system impulse response be $h(t) = e^{-at}u(t)$, $a > 0$. Compute $\Phi_{xx}(j\omega)$, $\Phi_{xy}(j\omega)$, and $\Phi_{yy}(j\omega)$ using the results of parts (a)–(d).

(f) Suppose that we are given the following Fourier transform of a function $\phi(t)$:

$$\Phi(j\omega) = \frac{\omega^2 + 100}{\omega^2 + 25}.$$

Find the impulse responses of *two* causal, stable LTI systems that have autocorrelation functions equal to $\phi(t)$. Which one of these has a causal, stable inverse?

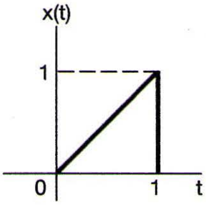


Figure P4.50

4.51. (a) Consider two LTI systems with impulse responses $h(t)$ and $g(t)$, respectively, and suppose that these systems are inverses of one another. Suppose also that the systems have frequency responses denoted by $H(j\omega)$ and $G(j\omega)$, respectively. What is the relationship between $H(j\omega)$ and $G(j\omega)$?

(b) Consider the continuous-time LTI system with frequency response

$$H(j\omega) = \begin{cases} 1, & 2 < |\omega| < 3 \\ 0, & \text{otherwise} \end{cases}.$$

(i) Is it possible to find an input $x(t)$ to this system such that the output is as depicted in Figure P4.50? If so, find $x(t)$. If not, explain why not.

(ii) Is this system invertible? Explain your answer.

(c) Consider an auditorium with an echo problem. As discussed in Problem 2.64, we can model the acoustics of the auditorium as an LTI system with an impulse response consisting of an impulse train, with the kth impulse in the train corresponding to the kth echo. Suppose that in this particular case the impulse response is

$$h(t) = \sum_{k=0}^{\infty} e^{-kT}\,\delta(t - kT),$$

where the factor e^{-kT} represents the attenuation of the kth echo.

In order to make a high-quality recording from the stage, the effect of the echoes must be removed by performing some processing of the sounds sensed by the recording equipment. In Problem 2.64, we used convolutional techniques to consider one example of the design of such a processor (for a different acoustic model). In the current problem, we will use frequency-domain techniques. Specifically, let $G(j\omega)$ denote the frequency response of the LTI system to be used to process the sensed acoustic signal. Choose $G(j\omega)$ so that the echoes are completely removed and the resulting signal is a faithful reproduction of the original stage sounds.

(d) Find the differential equation for the inverse of the system with impulse response

$$h(t) = 2\,\delta(t) + u_1(t).$$

(e) Consider the LTI system initially at rest and described by the differential equation

$$\frac{d^2y(t)}{dt^2} + 6\frac{dy(t)}{dt} + 9y(t) = \frac{d^2x(t)}{dt^2} + 3\frac{dx(t)}{dt} + 2x(t).$$

The inverse of this system is also initially at rest and described by a differential equation. Find the differential equation describing the inverse, and find the impulse responses $h(t)$ and $g(t)$ of the original system and its inverse.

4.52. Inverse systems frequently find application in problems involving imperfect measuring devices. For example, consider a device for measuring the temperature of a liquid. It is often reasonable to model such a device as an LTI system that, because of the response characteristics of the measuring element (e.g., the mercury in a thermometer), does not respond instantaneously to temperature changes. In particular, assume that the response of this device to a unit step in temperature is

$$s(t) = (1 - e^{-t/2})u(t). \qquad \text{(P4.52–1)}$$

(a) Design a compensatory system that, when provided with the output of the measuring device, produces an output equal to the instantaneous temperature of the liquid.

(b) One of the problems that often arises in using inverse systems as compensators for measuring devices is that gross inaccuracies in the indicated temperature may occur if the actual output of the measuring device produces errors due to small, erratic phenomena in the device. Since there always are such sources of error in real systems, one must take them into account. To illustrate this, consider a measuring device whose overall output can be modeled as the sum of the response of the measuring device characterized by eq. (P4.52–1) and an interfering "noise" signal $n(t)$. Such a model is depicted in Figure P4.52(a), where we have also included the inverse system of part (a), which now has as its input the *overall* output of the measuring device. Suppose that $n(t) = \sin\omega t$. What is the contribution of $n(t)$ to the output of the inverse system, and how does this output change as ω is increased?

(c) The issue raised in part (b) is an important one in many applications of LTI system analysis. Specifically, we are confronted with the fundamental tradeoff between the speed of response of the system and the ability of the system to attenuate high-frequency interference. In part (b) we saw that this tradeoff implied that, by attempting to speed up the response of a measuring device (by means of an inverse system), we produced a system that would also amplify

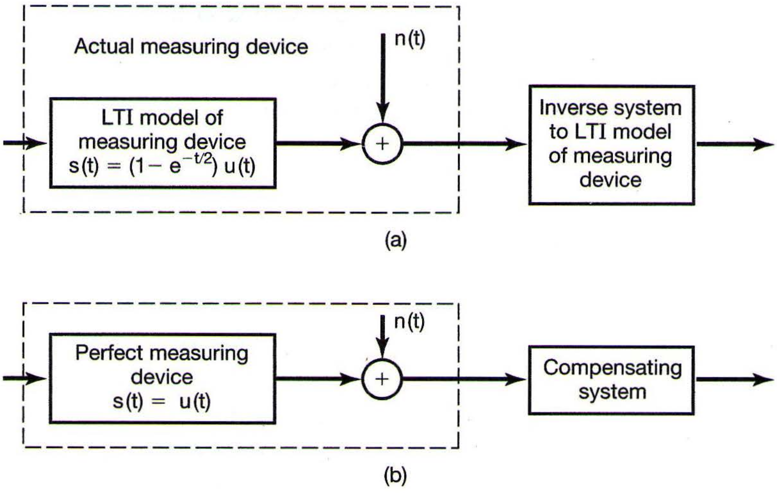

Actual measuring device
n(t)
LTI model of measuring device s(t) = (1 − e^{−t/2}) u(t)
+
Inverse system to LTI model of measuring device
(a)
n(t)
Perfect measuring device s(t) = u(t)
+
Compensating system
(b)

Figure P4.52

corrupting sinusoidal signals. To illustrate this concept further, consider a measuring device that responds instantaneously to changes in temperature, but that also is corrupted by noise. The response of such a system can be modeled, as depicted in Figure P4.52(b), as the sum of the response of a perfect measuring device and a corrupting signal $n(t)$. Suppose that we wish to design a compensatory system that will *slow down* the response to actual temperature variations, but also will attenuate the noise $n(t)$. Let the impulse response of this system be

$$h(t) = ae^{-at}u(t).$$

Choose a so that the overall system of Figure P4.52(b) responds as quickly as possible to a step change in temperature, subject to the constraint that the amplitude of the portion of the output due to the noise $n(t) = \sin 6t$ is no larger than 1/4.

4.53. As mentioned in the text, the techniques of Fourier analysis can be extended to signals having two independent variables. As their one-dimensional counterparts do in some applications, these techniques play an important role in other applications, such as image processing. In this problem, we introduce some of the elementary ideas of two-dimensional Fourier analysis.

Let $x(t_1, t_2)$ be a signal that depends upon two independent variables t_1 and t_2. The *two-dimensional Fourier transform* of $x(t_1, t_2)$ is defined as

$$X(j\omega_1, j\omega_2) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x(t_1, t_2)e^{-j(\omega_1 t_1 + \omega_2 t_2)}\,dt_1\,dt_2.$$

(a) Show that this double integral can be performed as two successive one-dimensional Fourier transforms, first in t_1 with t_2 regarded as fixed and then in t_2.

(b) Use the result of part (a) to determine the inverse transform—that is, an expression for $x(t_1, t_2)$ in terms of $X(j\omega_1, j\omega_2)$.

(c) Determine the two-dimensional Fourier transforms of the following signals:

(i) $x(t_1, t_2) = e^{-t_1+2t_2}u(t_1 - 1)u(2 - t_2)$

(ii) $x(t_1, t_2) = \begin{cases} e^{-|t_1|-|t_2|}, & \text{if } -1 < t_1 \leq 1 \text{ and } -1 \leq t_2 \leq 1 \\ 0, & \text{otherwise} \end{cases}$

(iii) $x(t_1, t_2) = \begin{cases} e^{-|t_1|-|t_2|}, & \text{if } 0 \leq t_1 \leq 1 \text{ or } 0 \leq t_2 \leq 1 \text{ (or both)} \\ 0, & \text{otherwise} \end{cases}$

(iv) $x(t_1, t_2)$ as depicted in Figure P4.53.

(v) $e^{-|t_1+t_2|-|t_1-t_2|}$

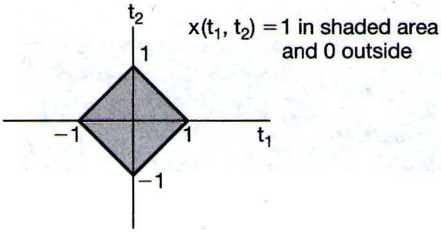


Figure P4.53

(d) Determine the signal $x(t_1, t_2)$ whose two-dimensional Fourier transform is

$$X(j\omega_1, j\omega_2) = \frac{2\pi}{4 + j\omega_1}\delta(\omega_2 - 2\omega_1).$$

(e) Let $x(t_1, t_2)$ and $h(t_1, t_2)$ be two signals with two-dimensional Fourier transforms $X(j\omega_1, j\omega_2)$ and $H(j\omega_1, j\omega_2)$, respectively. Determine the transforms of the following signals in terms of $X(j\omega_1, j\omega_2)$ and $H(j\omega_1, j\omega_2)$:

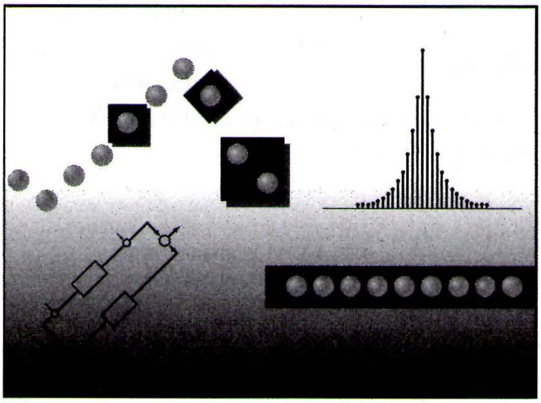
(i) $x(t_1 - T_1, t_2 - T_2)$

(ii) $x(at_1, bt_2)$

(iii) $y(t_1, t_2) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x(\tau_1, \tau_2)h(t_1 - \tau_1, t_2 - \tau_2)\,d\tau_1\,d\tau_2$

5

THE DISCRETE-TIME FOURIER TRANSFORM



5.0 INTRODUCTION

In Chapter 4, we introduced the continuous-time Fourier transform and developed the many characteristics of that transform which make the methods of Fourier analysis of such great value in analyzing and understanding the properties of continuous-time signals and systems. In the current chapter, we complete our development of the basic tools of Fourier analysis by introducing and examining the discrete-time Fourier transform.

In our discussion of Fourier series in Chapter 3, we saw that there are many similarities and strong parallels in analyzing continuous-time and discrete-time signals. However, there are also important differences. For example, as we saw in Section 3.6, the Fourier series representation of a discrete-time periodic signal is a *finite* series, as opposed to the infinite series representation required for continuous-time periodic signals. As we will see in this chapter, there are corresponding differences between continuous-time and discrete-time Fourier transforms.

In the remainder of the chapter, we take advantage of the similarities between continuous-time and discrete-time Fourier analysis by following a strategy essentially identical to that used in Chapter 4. In particular, we begin by extending the Fourier series description of periodic signals in order to develop a Fourier transform representation for discrete-time aperiodic signals, and we follow with an analysis of the properties and characteristics of the discrete-time Fourier transform that parallels that given in Chapter 4. By doing this, we not only will enhance our understanding of the basic concepts of Fourier analysis that are common to both continuous and discrete time, but also will contrast their differences in order to deepen our understanding of the distinct characteristics of each.

5.1 REPRESENTATION OF APERIODIC SIGNALS: THE DISCRETE-TIME FOURIER TRANSFORM

5.1.1 Development of the Discrete-Time Fourier Transform

In Section 4.1 [eq. (4.2) and Figure 4.2], we saw that the Fourier series coefficients for a continuous-time periodic square wave can be viewed as samples of an envelope function and that, as the period of the square wave increases, these samples become more and more finely spaced. This property suggested representing an aperiodic signal $x(t)$ by first constructing a periodic signal $\tilde{x}(t)$ that equaled $x(t)$ over one period. Then, as this period approached infinity, $\tilde{x}(t)$ was equal to $x(t)$ over larger and larger intervals of time, and the Fourier series representation for $\tilde{x}(t)$ converged to the Fourier transform representation for $x(t)$. In this section, we apply an analogous procedure to discrete-time signals in order to develop the Fourier transform representation for discrete-time aperiodic sequences.

Consider a general sequence $x[n]$ that is of finite duration. That is, for some integers N_1 and N_2, $x[n] = 0$ outside the range $-N_1 \le n \le N_2$. A signal of this type is illustrated in Figure 5.1(a). From this aperiodic signal, we can construct a periodic sequence $\tilde{x}[n]$ for which $x[n]$ is one period, as illustrated in Figure 5.1(b). As we choose the period N to be larger, $\tilde{x}[n]$ is identical to $x[n]$ over a longer interval, and as $N \to \infty$, $\tilde{x}[n] = x[n]$ for any finite value of n.

Let us now examine the Fourier series representation of $\tilde{x}[n]$. Specifically, from eqs. (3.94) and (3.95), we have

$$\tilde{x}[n] = \sum_{k=\langle N \rangle} a_k e^{jk(2\pi/N)n}, \tag{5.1}$$

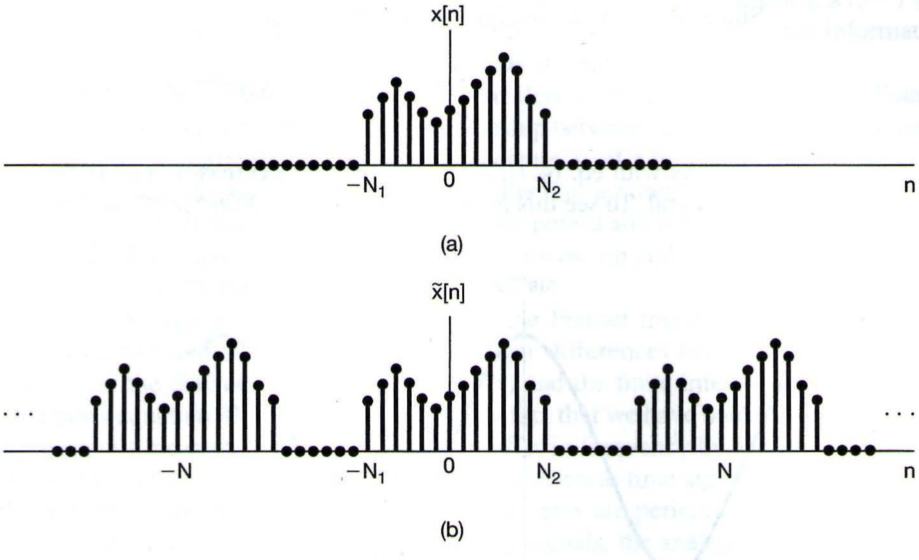


Figure 5.1 (a) Finite-duration signal $x[n]$; (b) periodic signal $\tilde{x}[n]$ constructed to be equal to $x[n]$ over one period.

$$a_k = \frac{1}{N} \sum_{n=\langle N \rangle} \tilde{x}[n] e^{-jk(2\pi/N)n}. \tag{5.2}$$

Since $x[n] = \tilde{x}[n]$ over a period that includes the interval $-N_1 \le n \le N_2$, it is convenient to choose the interval of summation in eq. (5.2) to include this interval, so that $\tilde{x}[n]$ can be replaced by $x[n]$ in the summation. Therefore,

$$a_k = \frac{1}{N} \sum_{n=-N_1}^{N_2} x[n] e^{-jk(2\pi/N)n} = \frac{1}{N} \sum_{n=-\infty}^{+\infty} x[n] e^{-jk(2\pi/N)n}, \tag{5.3}$$

where in the second equality in eq. (5.3) we have used the fact that $x[n]$ is zero outside the interval $-N_1 \le n \le N_2$. Defining the function

$$X(e^{j\omega}) = \sum_{n=-\infty}^{+\infty} x[n] e^{-j\omega n}, \tag{5.4}$$

we see that the coefficients a_k are proportional to samples of $X(e^{j\omega})$, i.e.,

$$a_k = \frac{1}{N} X(e^{jk\omega_0}), \tag{5.5}$$

where $\omega_0 = 2\pi/N$ is the spacing of the samples in the frequency domain. Combining eqs. (5.1) and (5.5) yields

$$\tilde{x}[n] = \sum_{k=\langle N \rangle} \frac{1}{N} X(e^{jk\omega_0}) e^{jk\omega_0 n}. \tag{5.6}$$

Since $\omega_0 = 2\pi/N$, or equivalently, $1/N = \omega_0/2\pi$, eq. (5.6) can be rewritten as

$$\tilde{x}[n] = \frac{1}{2\pi} \sum_{k=\langle N \rangle} X(e^{jk\omega_0}) e^{jk\omega_0 n} \omega_0. \tag{5.7}$$

As with eq. (4.7), as N increases ω_0 decreases, and as $N \to \infty$ eq. (5.7) passes to an integral. To see this more clearly, consider $X(e^{j\omega})e^{j\omega n}$ as sketched in Figure 5.2. From

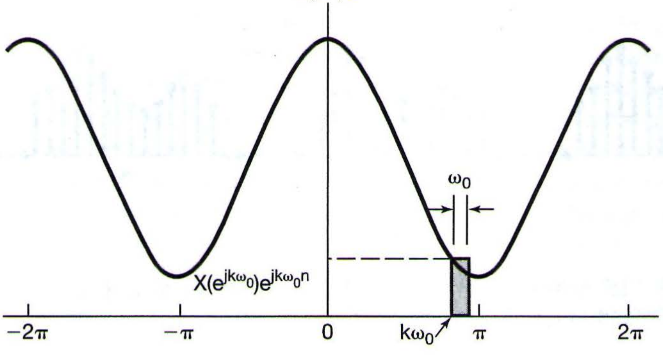

$X(e^{j\omega})e^{j\omega n}$
ω_0
$X(e^{jk\omega_0})e^{jk\omega_0 n}$
-2π
$-\pi$
0
$k\omega_0$
π
2π

Figure 5.2 Graphical interpretation of eq. (5.7).

eq. (5.4), $X(e^{j\omega})$ is seen to be periodic in ω with period 2π, and so is $e^{j\omega n}$. Thus, the product $X(e^{j\omega})e^{j\omega n}$ will also be periodic. As depicted in the figure, each term in the summation in eq. (5.7) represents the area of a rectangle of height $X(e^{jk\omega_0})e^{j\omega_0 n}$ and width ω_0. As $\omega_0 \to 0$, the summation becomes an integral. Furthermore, since the summation is carried out over N consecutive intervals of width $\omega_0 = 2\pi/N$, the total interval of integration will always have a width of 2π. Therefore, as $N \to \infty$, $\tilde{x}[n] = x[n]$, and eq. (5.7) becomes

$$x[n] = \frac{1}{2\pi}\int_{2\pi} X(e^{j\omega})e^{j\omega n}\,d\omega,$$

where, since $X(e^{j\omega})e^{j\omega n}$ is periodic with period 2π, the interval of integration can be taken as *any* interval of length 2π. Thus, we have the following pair of equations:

$$\boxed{\begin{aligned} x[n] &= \frac{1}{2\pi}\int_{2\pi} X(e^{j\omega})e^{j\omega n}\,d\omega, && (5.8)\\ X(e^{j\omega}) &= \sum_{n=-\infty}^{+\infty} x[n]e^{-j\omega n}. && (5.9)\end{aligned}}$$

Equations (5.8) and (5.9) are the discrete-time counterparts of eqs. (4.8) and (4.9). The function $X(e^{j\omega})$ is referred to as the *discrete-time Fourier transform* and the pair of equations as the *discrete-time Fourier transform pair.* Equation (5.8) is the *synthesis equation,* eq. (5.9) the *analysis equation.* Our derivation of these equations indicates how an aperiodic sequence can be thought of as a linear combination of complex exponentials. In particular, the synthesis equation is in effect a representation of $x[n]$ as a linear combination of complex exponentials infinitesimally close in frequency and with amplitudes $X(e^{j\omega})(d\omega/2\pi)$. For this reason, as in continuous time, the Fourier transform $X(e^{j\omega})$ will often be referred to as the *spectrum* of $x[n]$, because it provides us with the information on how $x[n]$ is composed of complex exponentials at different frequencies.

Note also that, as in continuous time, our derivation of the discrete-time Fourier transform provides us with an important relationship between discrete-time Fourier series and transforms. In particular, the Fourier coefficients a_k of a periodic signal $\tilde{x}[n]$ can be expressed in terms of equally spaced *samples* of the Fourier transform of a finite-duration, aperiodic signal $x[n]$ that is equal to $\tilde{x}[n]$ over one period and is zero otherwise. This fact is of considerable importance in practical signal processing and Fourier analysis, and we look at it further in Problem 5.41.

As our derivation indicates, the discrete-time Fourier transform shares many similarities with the continuous-time case. The major differences between the two are the periodicity of the discrete-time transform $X(e^{j\omega})$ and the finite interval of integration in the synthesis equation. Both of these stem from a fact that we have noted several times before: Discrete-time complex exponentials that differ in frequency by a multiple of 2π are identical. In Section 3.6 we saw that, for periodic discrete-time signals, the implications of this statement are that the Fourier series coefficients are periodic and that the Fourier series representation is a finite sum. For aperiodic signals, the analogous implications are that $X(e^{j\omega})$ is periodic (with period 2π) and that the synthesis equation involves an integration only over a frequency interval that produces distinct complex exponentials (i.e., any interval of length 2π). In Section 1.3.3, we noted one further consequence of the periodicity of $e^{j\omega n}$ as a function of ω: $\omega = 0$ and $\omega = 2\pi$ yield the same signal. Signals at frequencies near these values or any other even multiple of π are slowly varying and therefore are all appropriately thought of as low-frequency signals. Similarly, the high frequencies in discrete time are the values of ω near odd multiples of π. Thus, the signal $x_1[n]$ shown in Figure 5.3(a) with Fourier transform depicted in Figure 5.3(b) varies more slowly than the signal $x_2[n]$ in Figure 5.3(c) whose transform is shown in Figure 5.3(d).

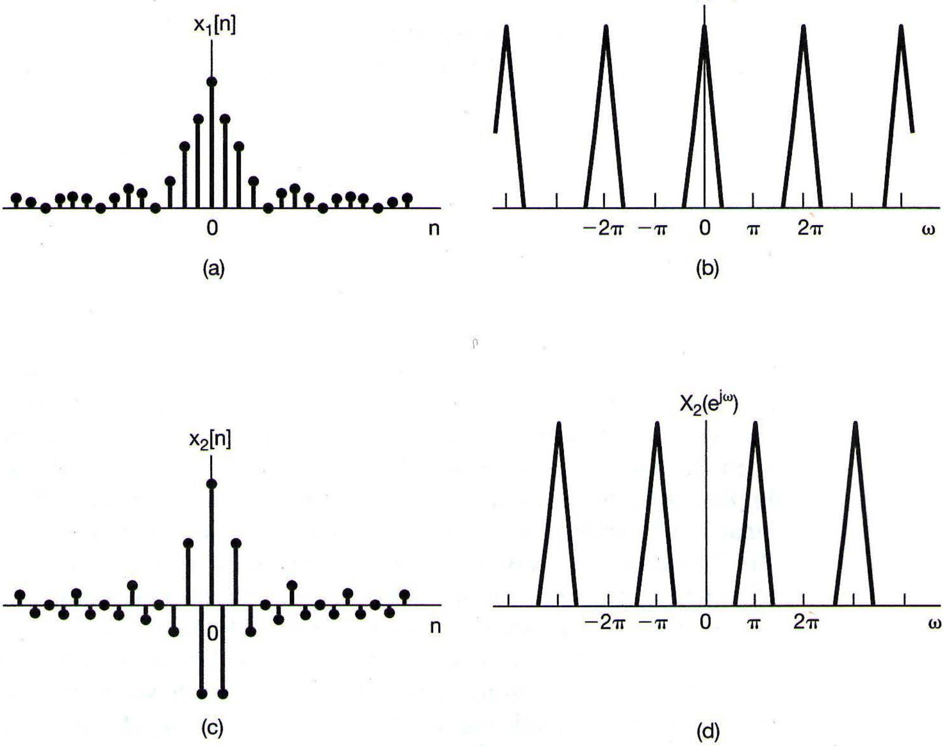


Figure 5.3 (a) Discrete-time signal $x_1[n]$. (b) Fourier transform of $x_1[n]$. Note that $X_1(e^{j\omega})$ is concentrated near $\omega = 0, \pm 2\pi, \pm 4\pi, \ldots$. (c) Discrete-time signal $x_2[n]$. (d) Fourier transform of $x_2[n]$. Note that $X_2(e^{j\omega})$ is concentrated near $\omega = \pm\pi, \pm 3\pi, \ldots$.

5.1.2 Examples of Discrete-Time Fourier Transforms

To illustrate the discrete-time Fourier transform, let us consider several examples.

Example 5.1

Consider the signal

$$x[n] = a^n u[n], \qquad |a| < 1.$$

In this case,

$$X(e^{j\omega}) = \sum_{n=-\infty}^{+\infty} a^n u[n] e^{-j\omega n}$$

$$= \sum_{n=0}^{\infty} (ae^{-j\omega})^n = \frac{1}{1 - ae^{-j\omega}}.$$

The magnitude and phase of $X(e^{j\omega})$ are shown in Figure 5.4(a) for $a > 0$ and in Figure 5.4(b) for $a < 0$. Note that all of these functions are periodic in ω with period 2π.

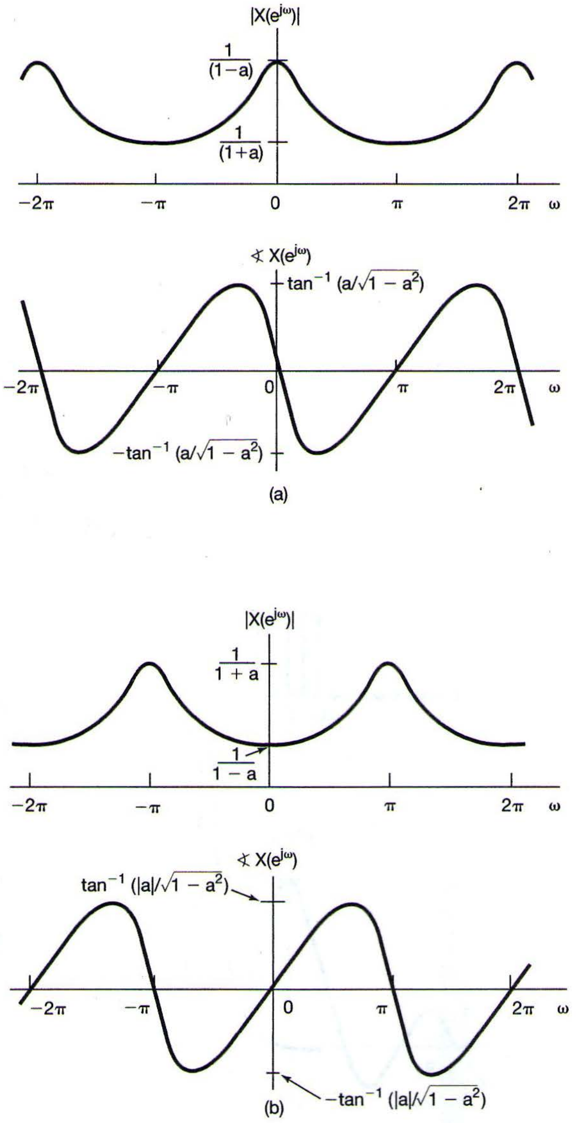


Figure 5.4 Magnitude and phase of the Fourier transform of Example 5.1 for (a) $a > 0$ and (b) $a < 0$.

Example 5.2

Let

$$x[n] = a^{|n|}, \qquad |a| < 1.$$

This signal is sketched for $0 < a < 1$ in Figure 5.5(a). Its Fourier transform is obtained from eq. (5.9):

$$\begin{aligned} X(e^{j\omega}) &= \sum_{n=-\infty}^{+\infty} a^{|n|} e^{-j\omega n} \\ &= \sum_{n=0}^{\infty} a^{n} e^{-j\omega n} + \sum_{n=-\infty}^{-1} a^{-n} e^{-j\omega n}. \end{aligned}$$

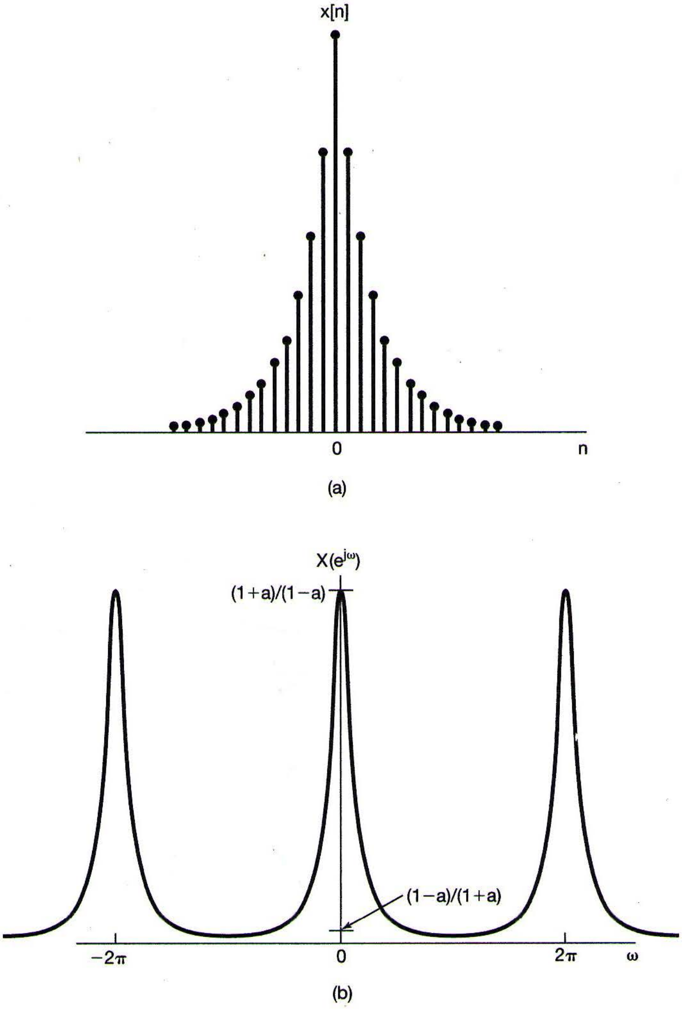


Figure 5.5 (a) Signal $x[n] = a^{|n|}$ of Example 5.2 and (b) its Fourier transform ($0 < a < 1$).

Making the substitution of variables $m = -n$ in the second summation, we obtain

$$X(e^{j\omega}) = \sum_{n=0}^{\infty}(ae^{-j\omega})^n + \sum_{m=1}^{\infty}(ae^{j\omega})^m.$$

Both of these summations are infinite geometric series that we can evaluate in closed form, yielding

$$\begin{aligned} X(e^{j\omega}) &= \frac{1}{1-ae^{-j\omega}} + \frac{ae^{j\omega}}{1-ae^{j\omega}} \\ &= \frac{1-a^2}{1-2a\cos\omega + a^2}. \end{aligned}$$

In this case, $X(e^{j\omega})$ is real and is illustrated in Figure 5.5(b), again for $0 < a < 1$.

Example 5.3

Consider the rectangular pulse

$$x[n] = \begin{cases} 1, & |n| \le N_1 \\ 0, & |n| > N_1 \end{cases}, \tag{5.10}$$

which is illustrated in Figure 5.6(a) for $N_1 = 2$. In this case,

$$X(e^{j\omega}) = \sum_{n=-N_1}^{N_1} e^{-j\omega n}. \tag{5.11}$$

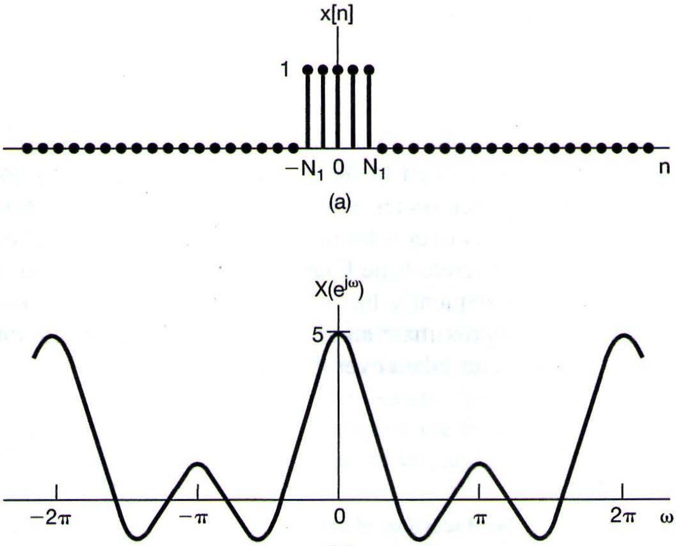


Figure 5.6 (a) Rectangular pulse signal of Example 5.3 for $N_1 = 2$ and (b) its Fourier transform.

Using calculations similar to those employed in obtaining eq. (3.104) in Example 3.12, we can write

$$X(e^{j\omega}) = \frac{\sin \omega\left(N_1 + \frac{1}{2}\right)}{\sin(\omega/2)}. \tag{5.12}$$

This Fourier transform is sketched in Figure 5.6(b) for $N_1 = 2$. The function in eq. (5.12) is the discrete-time counterpart of the sinc function, which appears in the Fourier transform of the continuous-time rectangular pulse (see Example 4.4). An important difference between these two functions is that the function in eq. (5.12) is periodic with period 2π, whereas the sinc function is aperiodic.

5.1.3 Convergence Issues Associated with the Discrete-Time Fourier Transform

Although the argument we used to derive the discrete-time Fourier transform in Section 5.1.1 was constructed assuming that $x[n]$ was of arbitrary but finite duration, eqs. (5.8) and (5.9) remain valid for an extremely broad class of signals with infinite duration (such as the signals in Examples 5.1 and 5.2). In this case, however, we again must consider the question of convergence of the infinite summation in the analysis equation (5.9). The conditions on $x[n]$ that guarantee the convergence of this sum are direct counterparts of the convergence conditions for the continuous-time Fourier transform.[1] Specifically, eq. (5.9) will converge either if $x[n]$ is absolutely summable, that is,

$$\sum_{n=-\infty}^{+\infty} |x[n]| < \infty, \tag{5.13}$$

or if the sequence has finite energy, that is,

$$\sum_{n=-\infty}^{+\infty} |x[n]|^2 < \infty. \tag{5.14}$$

In contrast to the situation for the analysis equation (5.9), there are generally no convergence issues associated with the synthesis equation (5.8), since the integral in this equation is over a finite interval of integration. This is very much the same situation as for the discrete-time Fourier series synthesis equation (3.94), which involves a finite sum and consequently has no issues of convergence associated with it either. In particular, if we approximate an aperiodic signal $x[n]$ by an integral of complex exponentials with frequencies taken over the interval $|\omega| \leq W$, i.e.,

$$\hat{x}[n] = \frac{1}{2\pi}\int_{-W}^{W} X(e^{j\omega})e^{j\omega n}\, d\omega, \tag{5.15}$$

---

[1] For discussions of the convergence issues associated with the discrete-time Fourier transform, see A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing* (Englewood Cliffs, NJ: Prentice-Hall, Inc., 1989), and L. R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing* (Englewood Cliffs, NJ: Prentice-Hall, Inc., 1975).

then $\hat{x}[n] = x[n]$ for $W = \pi$. Thus, much as in Figure 3.18, we would expect not to see any behavior like the Gibbs phenomenon in evaluating the discrete-time Fourier transform synthesis equation. This is illustrated in the following example.

Example 5.4

Let $x[n]$ be the unit impulse; that is,

$$x[n] = \delta[n].$$

In this case the analysis equation (5.9) is easily evaluated, yielding

$$X(e^{j\omega}) = 1.$$

In other words, just as in continuous time, the unit impulse has a Fourier transform representation consisting of equal contributions at all frequencies. If we then apply eq. (5.15) to this example, we obtain

$$\hat{x}[n] = \frac{1}{2\pi}\int_{-W}^{W} e^{j\omega n}\,d\omega = \frac{\sin Wn}{\pi n}. \tag{5.16}$$

This is plotted in Figure 5.7 for several values of W. As can be seen, the frequency of the oscillations in the approximation increases as W is increased, which is similar to what we observed in the continuous-time case. On the other hand, in contrast to the continuous-time case, the amplitude of these oscillations decreases relative to the magnitude of $\hat{x}[0]$ as W is increased, and the oscillations disappear entirely for $W = \pi$.

5.2 THE FOURIER TRANSFORM FOR PERIODIC SIGNALS

As in the continuous-time case, discrete-time periodic signals can be incorporated within the framework of the discrete-time Fourier transform by interpreting the transform of a periodic signal as an impulse train in the frequency domain. To derive the form of this representation, consider the signal

$$x[n] = e^{j\omega_0 n}. \tag{5.17}$$

In continuous time, we saw that the Fourier transform of $e^{j\omega_0 t}$ can be interpreted as an impulse at $\omega = \omega_0$. Therefore, we might expect the same type of transform to result for the discrete-time signal of eq. (5.17). However, the discrete-time Fourier transform must be periodic in ω with period 2π. This then suggests that the Fourier transform of $x[n]$ in eq. (5.17) should have impulses at $\omega_0, \omega_0 \pm 2\pi, \omega_0 \pm 4\pi$, and so on. In fact, the Fourier transform of $x[n]$ is the impulse train

$$X(e^{j\omega}) = \sum_{l=-\infty}^{+\infty} 2\pi\,\delta(\omega - \omega_0 - 2\pi l), \tag{5.18}$$

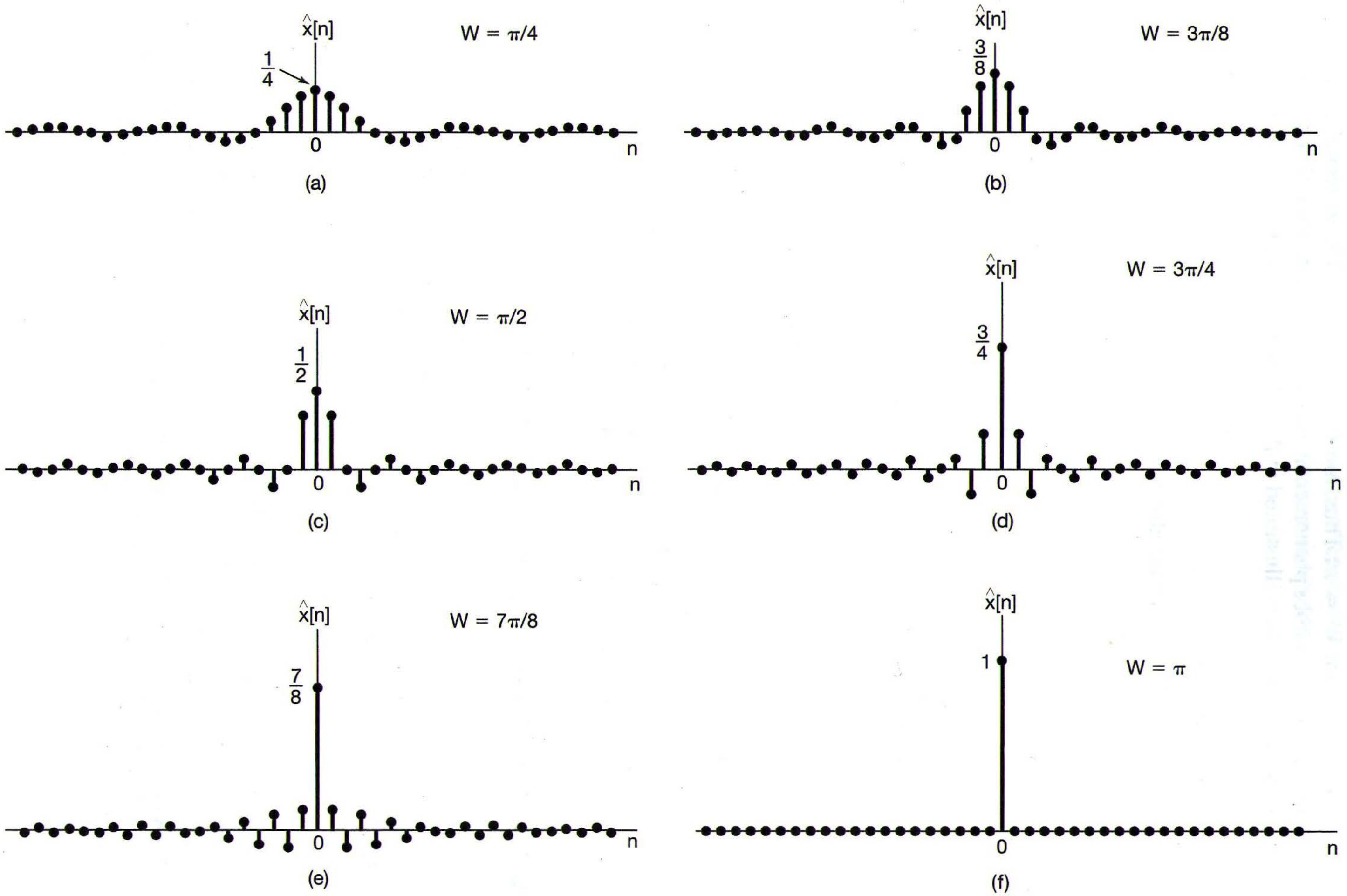


Figure 5.7 Approximation to the unit sample obtained as in eq. (5.16) using complex exponentials with frequencies $|\omega| \leq W$: (a) $W = \pi/4$; (b) $W = 3\pi/8$; (c) $W = \pi/2$; (d) $W = 3\pi/4$; (e) $W = 7\pi/8$; (f) $W = \pi$. Note that for $W = \pi$, $\hat{x}[n] = \delta[n]$.

which is illustrated in Figure 5.8. In order to check the validity of this expression, we must evaluate its inverse transform. Substituting eq. (5.18) into the synthesis equation (5.8), we find that

$$\frac{1}{2\pi}\int_{2\pi} X(e^{j\omega})e^{j\omega n}\,d\omega = \frac{1}{2\pi}\int_{2\pi} \sum_{l=-\infty}^{+\infty} 2\pi\,\delta(\omega-\omega_0-2\pi l)e^{j\omega n}\,d\omega.$$

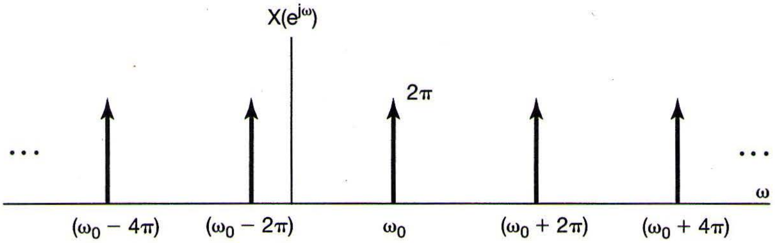


Figure 5.8 Fourier transform of $x[n] = e^{j\omega_0 n}$.

Note that any interval of length 2π includes exactly one impulse in the summation given in eq. (5.18). Therefore, if the interval of integration chosen includes the impulse located at $\omega_0 + 2\pi r$, then

$$\frac{1}{2\pi}\int_{2\pi} X(e^{j\omega})e^{j\omega n}\,d\omega = e^{j(\omega_0+2\pi r)n} = e^{j\omega_0 n}.$$

Now consider a periodic sequence $x[n]$ with period N and with the Fourier series representation

$$x[n] = \sum_{k=\langle N\rangle} a_k e^{jk(2\pi/N)n}. \tag{5.19}$$

In this case, the Fourier transform is

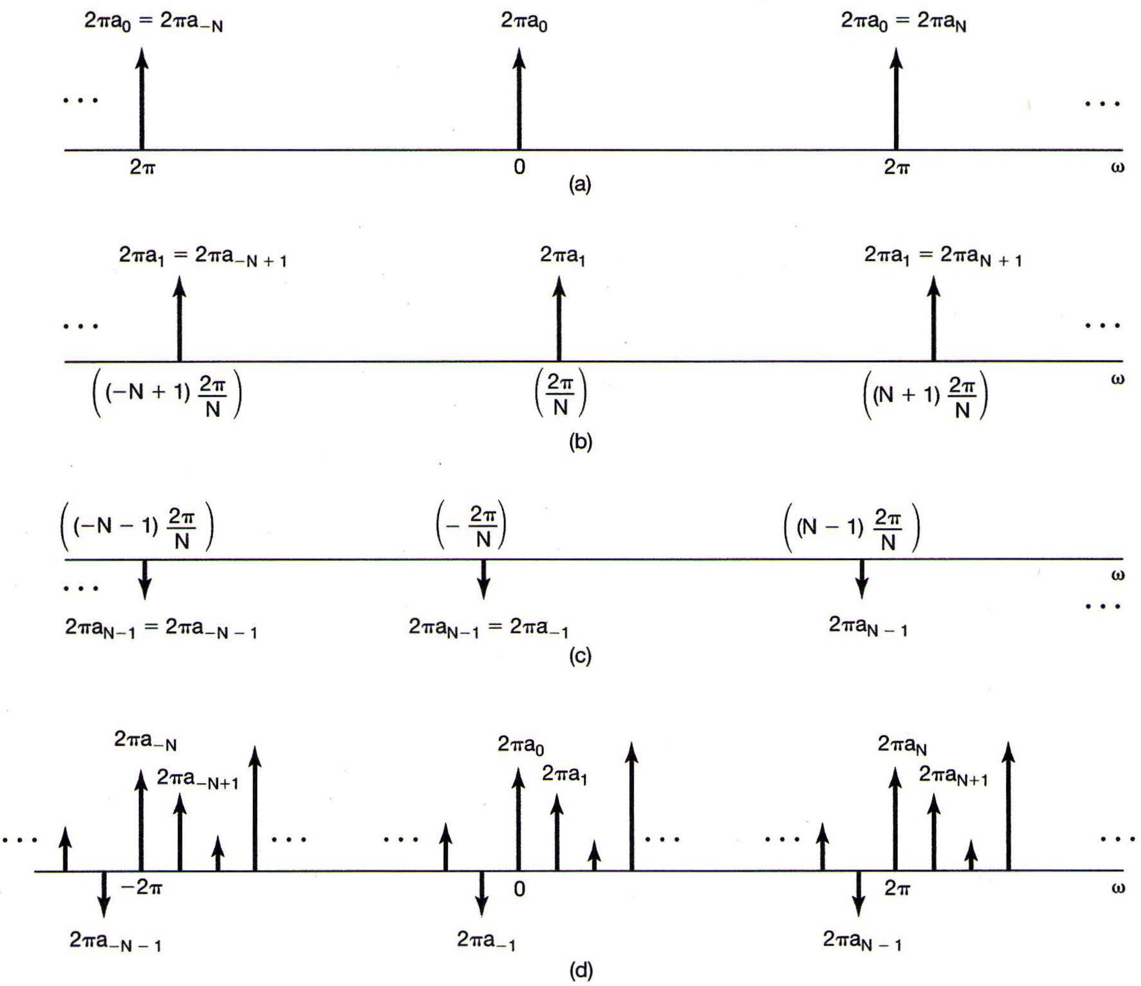
$$X(e^{j\omega}) = \sum_{k=-\infty}^{+\infty} 2\pi a_k\,\delta\left(\omega - \frac{2\pi k}{N}\right), \tag{5.20}$$

so that the Fourier transform of a periodic signal can be directly constructed from its Fourier coefficients.

To verify that eq. (5.20) is in fact correct, note that $x[n]$ in eq. (5.19) is a linear combination of signals of the form in eq. (5.17), and thus the Fourier transform of $x[n]$ must be a linear combination of transforms of the form of eq. (5.18). In particular, suppose that we choose the interval of summation in eq. (5.19) as $k = 0, 1, \ldots, N-1$, so that

$$\begin{aligned} x[n] = {} & a_0 + a_1 e^{j(2\pi/N)n} + a_2 e^{j2(2\pi/N)n} \\ & + \cdots + a_{N-1} e^{j(N-1)(2\pi/N)n}. \end{aligned} \tag{5.21}$$

Thus, $x[n]$ is a linear combination of signals, as in eq. (5.17), with $\omega_0 = 0, 2\pi/N, 4\pi/N, \ldots, (N-1)2\pi/N$. The resulting Fourier transform is illustrated in Figure 5.9. In Figure 5.9(a), we have depicted the Fourier transform of the first term on the right-hand side of eq. (5.21): The Fourier transform of the constant signal $a_0 = a_0 e^{j0 \cdot n}$ is a periodic impulse train, as in eq. (5.18), with $\omega_0 = 0$ and a scaling of $2\pi a_0$ on each of the impulses. Moreover, from Chapter 4 we know that the Fourier series coefficients a_k are periodic with period N, so that $2\pi a_0 = 2\pi a_N = 2\pi a_{-N}$. In Figure 5.9(b) we have illustrated the Fourier transform of the second term in eq. (5.21), where we have again used eq. (5.18),



$2\pi a_0 = 2\pi a_{-N}$
$2\pi a_0$
$2\pi a_0 = 2\pi a_N$
2π
0
2π
ω
(a)
$2\pi a_1 = 2\pi a_{-N+1}$
$2\pi a_1$
$2\pi a_1 = 2\pi a_{N+1}$
$\left((-N+1)\frac{2\pi}{N}\right)$
$\left(\frac{2\pi}{N}\right)$
$\left((N+1)\frac{2\pi}{N}\right)$
ω
(b)
$\left((-N-1)\frac{2\pi}{N}\right)$
$\left(-\frac{2\pi}{N}\right)$
$\left((N-1)\frac{2\pi}{N}\right)$
ω
$2\pi a_{N-1} = 2\pi a_{-N-1}$
$2\pi a_{N-1} = 2\pi a_{-1}$
$2\pi a_{N-1}$
(c)
$2\pi a_{-N}$
$2\pi a_{-N+1}$
$2\pi a_0$
$2\pi a_1$
$2\pi a_N$
$2\pi a_{N+1}$
-2π
0
2π
ω
$2\pi a_{-N-1}$
$2\pi a_{-1}$
$2\pi a_{N-1}$
(d)

Figure 5.9 Fourier transform of a discrete-time periodic signal: (a) Fourier transform of the first term on the right-hand side of eq. (5.21); (b) Fourier transform of the second term in eq. (5.21); (c) Fourier transform of the last term in eq. (5.21); (d) Fourier transform of $x[n]$ in eq (5.21).

in this case for $a_1 e^{j(2\pi/N)n}$, and the fact that $2\pi a_1 = 2\pi a_{N+1} = 2\pi a_{-N+1}$. Similarly, Figure 5.9(c) depicts the final term. Finally, Figure 5.9(d) depicts the entire expression for $X(e^{j\omega})$. Note that because of the periodicity of the a_k, $X(e^{j\omega})$ can be interpreted as a train of impulses occurring at multiples of the fundamental frequency $2\pi/N$, with the area of the impulse located at $\omega = 2\pi k/N$ being $2\pi a_k$, which is exactly what is stated in eq. (5.20).

Example 5.5

Consider the periodic signal

$$x[n] = \cos\omega_0 n = \frac{1}{2}e^{j\omega_0 n} + \frac{1}{2}e^{-j\omega_0 n}, \qquad \text{with} \quad \omega_0 = \frac{2\pi}{5}. \tag{5.22}$$

From eq. (5.18), we can immediately write

$$X(e^{j\omega}) = \sum_{l=-\infty}^{+\infty} \pi\,\delta\left(\omega - \frac{2\pi}{5} - 2\pi l\right) + \sum_{l=-\infty}^{+\infty} \pi\,\delta\left(\omega + \frac{2\pi}{5} - 2\pi l\right). \tag{5.23}$$

That is,

$$X(e^{j\omega}) = \pi\,\delta\left(\omega - \frac{2\pi}{5}\right) + \pi\,\delta\left(\omega + \frac{2\pi}{5}\right), \qquad -\pi \le \omega < \pi, \tag{5.24}$$

and $X(e^{j\omega})$ repeats periodically with a period of 2π, as illustrated in Figure 5.10.

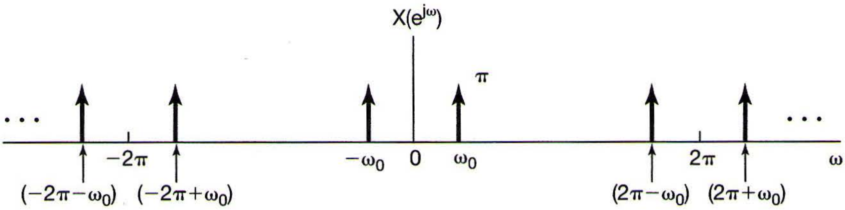


Figure 5.10 Discrete-time Fourier transform of $x[n] = \cos\omega_0 n$.

Example 5.6

The discrete-time counterpart of the periodic impulse train of Example 4.8 is the sequence

$$x[n] = \sum_{k=-\infty}^{+\infty} \delta[n - kN], \tag{5.25}$$

as sketched in Figure 5.11(a). The Fourier series coefficients for this signal can be calculated directly from eq. (3.95):

$$a_k = \frac{1}{N}\sum_{n=\langle N\rangle} x[n]e^{-jk(2\pi/N)n}.$$

Choosing the interval of summation as $0 \leq n \leq N-1$, we have

$$a_k = \frac{1}{N}. \tag{5.26}$$

Using eqs. (5.26) and (5.20), we can then represent the Fourier transform of the signal as

$$X(e^{j\omega}) = \frac{2\pi}{N} \sum_{k=-\infty}^{+\infty} \delta\left(\omega - \frac{2\pi k}{N}\right), \tag{5.27}$$

which is illustrated in Figure 5.11(b).

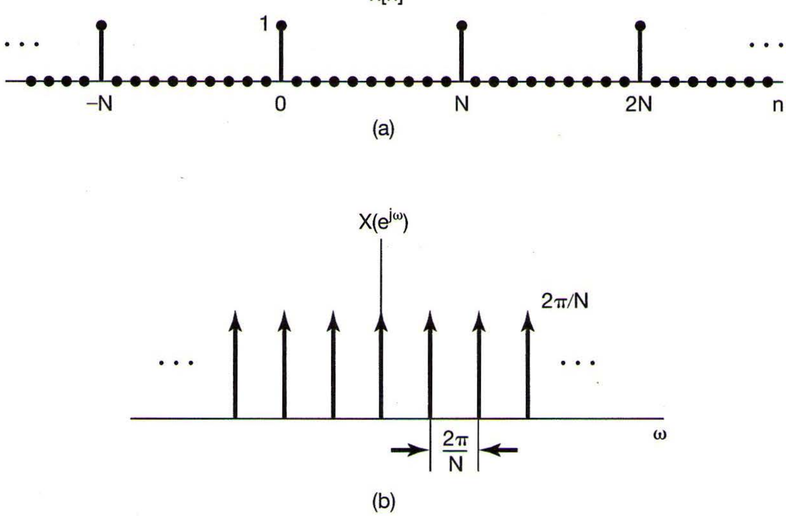


Figure 5.11 (a) Discrete-time periodic impulse train; (b) its Fourier transform.

5.3 PROPERTIES OF THE DISCRETE-TIME FOURIER TRANSFORM

As with the continuous-time Fourier transform, a variety of properties of the discrete-time Fourier transform provide further insight into the transform and, in addition, are often useful in reducing the complexity in the evaluation of transforms and inverse transforms. In this and the following two sections we consider these properties, and in Table 5.1 we present a concise summary of them. By comparing this table with Table 4.1, we can get a clear picture of some of the similarities and differences between continuous-time and discrete-time Fourier transform properties. When the derivation or interpretation of a discrete-time Fourier transform property is essentially identical to its continuous-time counterpart, we will simply state the property. Also, because of the close relationship between the Fourier series and the Fourier transform, many of the transform properties translate directly into corresponding properties for the discrete-time Fourier series, which we summarized in Table 3.2 and briefly discussed in Section 3.7.

In the following discussions, it will be convenient to adopt notation similar to that used in Section 4.3 to indicate the pairing of a signal and its transform. That is,

$$\begin{aligned} X(e^{j\omega}) &= \mathfrak{F}\{x[n]\}, \\ x[n] &= \mathfrak{F}^{-1}\{X(e^{j\omega})\}, \\ x[n] &\overset{\mathfrak{F}}{\longleftrightarrow} X(e^{j\omega}). \end{aligned}$$

5.3.1 Periodicity of the Discrete-Time Fourier Transform

As we discussed in Section 5.1, the discrete-time Fourier transform is *always* periodic in ω with period 2π; i.e.,

$$\boxed{X(e^{j(\omega+2\pi)}) = X(e^{j\omega}).} \tag{5.28}$$

This is in contrast to the continuous-time Fourier transform, which in general is not periodic.

5.3.2 Linearity of the Fourier Transform

If

$$x_1[n] \overset{\mathfrak{F}}{\longleftrightarrow} X_1(e^{j\omega})$$

and

$$x_2[n] \overset{\mathfrak{F}}{\longleftrightarrow} X_2(e^{j\omega}),$$

then

$$\boxed{ax_1[n] + bx_2[n] \overset{\mathfrak{F}}{\longleftrightarrow} aX_1(e^{j\omega}) + bX_2(e^{j\omega}).} \tag{5.29}$$

5.3.3 Time Shifting and Frequency Shifting

If

$$x[n] \overset{\mathfrak{F}}{\longleftrightarrow} X(e^{j\omega}),$$

then

$$\boxed{x[n - n_0] \overset{\mathfrak{F}}{\longleftrightarrow} e^{-j\omega n_0}X(e^{j\omega})} \tag{5.30}$$

and

$$\boxed{e^{j\omega_0 n}x[n] \overset{\mathfrak{F}}{\longleftrightarrow} X(e^{j(\omega-\omega_0)}).} \tag{5.31}$$

Equation (5.30) can be obtained by direct substitution of $x[n-n_0]$ into the analysis equation (5.9), while eq. (5.31) is derived by substituting $X(e^{j(\omega-\omega_0)})$ into the synthesis equation (5.8).

As a consequence of the periodicity and frequency-shifting properties of the discrete-time Fourier transform, there exists a special relationship between ideal lowpass and ideal highpass discrete-time filters. This is illustrated in the next example.

Example 5.7

In Figure 5.12(a) we have depicted the frequency response $H_{\text{lp}}(e^{j\omega})$ of a lowpass filter with cutoff frequency ω_c, while in Figure 5.12(b) we have displayed $H_{\text{lp}}(e^{j(\omega-\pi)})$—that is, the frequency response $H_{\text{lp}}(e^{j\omega})$ shifted by one-half period, i.e., by π. Since high frequencies in discrete time are concentrated near π (and other odd multiples of π), the filter in Figure 5.12(b) is an ideal highpass filter with cutoff frequency $\pi - \omega_c$. That is,

$$H_{\text{hp}}(e^{j\omega}) = H_{\text{lp}}(e^{j(\omega-\pi)}). \tag{5.32}$$

As we can see from eq. (3.122), and as we will discuss again in Section 5.4, the frequency response of an LTI system is the Fourier transform of the impulse response of the system. Thus, if $h_{\text{lp}}[n]$ and $h_{\text{hp}}[n]$ respectively denote the impulse responses of

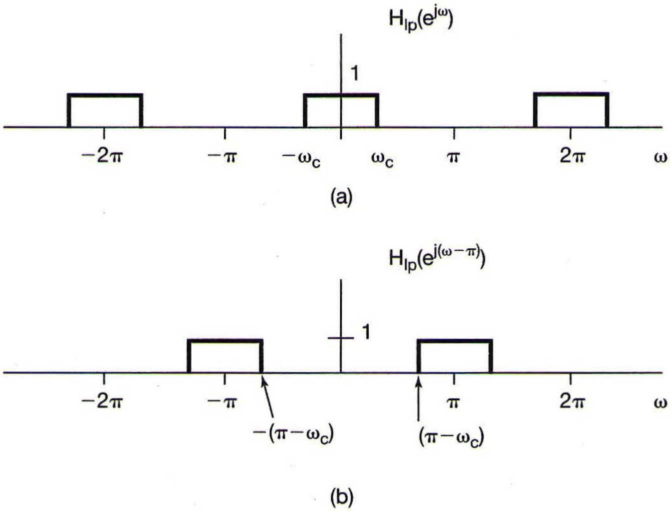


Figure 5.12 (a) Frequency response of a lowpass filter; (b) frequency response of a highpass filter obtained by shifting the frequency response in (a) by $\omega = \pi$ corresponding to one-half period.

The lowpass and highpass filters in Figure 5.12, then eq. (5.32) and the frequency-shifting property imply that

$$h_{\mathrm{hp}}[n] = e^{j\pi n} h_{\mathrm{lp}}[n] \tag{5.33}$$

$$= (-1)^n h_{\mathrm{lp}}[n]. \tag{5.34}$$

5.3.4 Conjugation and Conjugate Symmetry

If

$$x[n] \overset{\mathfrak{F}}{\longleftrightarrow} X(e^{j\omega}),$$

then

$$\boxed{x^*[n] \overset{\mathfrak{F}}{\longleftrightarrow} X^*(e^{-j\omega}).} \tag{5.35}$$

Also, if $x[n]$ is real valued, its transform $X(e^{j\omega})$ is conjugate symmetric. That is,

$$\boxed{X(e^{j\omega}) = X^*(e^{-j\omega}) \quad [x[n]\text{real}].} \tag{5.36}$$

From this, it follows that $\mathcal{R}e\{X(e^{j\omega})\}$ is an even function of ω and $\mathcal{I}m\{X(e^{j\omega})\}$ is an odd function of ω. Similarly, the magnitude of $X(e^{j\omega})$ is an even function and the phase angle is an odd function. Furthermore,

$$\mathcal{E}v\{x[n]\} \overset{\mathfrak{F}}{\longleftrightarrow} \mathcal{R}e\{X(e^{j\omega})\}$$

and

$$\mathcal{O}d\{x[n]\} \overset{\mathfrak{F}}{\longleftrightarrow} j\mathcal{I}m\{X(e^{j\omega})\},$$

where $\mathcal{E}v$ and $\mathcal{O}d$ denote the even and odd parts, respectively, of $x[n]$. For example, if $x[n]$ is real and even, its Fourier transform is also real and even. Example 5.2 illustrates this symmetry for $x[n] = a^{|n|}$.

5.3.5 Differencing and Accumulation

In this subsection, we consider the discrete-time counterpart of integration—that is, accumulation—and its inverse, first differencing. Let $x[n]$ be a signal with Fourier transform $X(e^{j\omega})$. Then, from the linearity and time-shifting properties, the Fourier transform pair for the first-difference signal $x[n] - x[n-1]$ is given by

$$\boxed{x[n] - x[n-1] \overset{\mathfrak{F}}{\longleftrightarrow} (1 - e^{-j\omega})X(e^{j\omega}).} \tag{5.37}$$

Next, consider the signal

$$y[n] = \sum_{m=-\infty}^{n} x[m]. \tag{5.38}$$

Since $y[n] - y[n-1] = x[n]$, we might conclude that the transform of $y[n]$ should be related to the transform of $x[n]$ by division by $(1 - e^{-j\omega})$. This is partly correct, but as with the continuous-time integration property given by eq. (4.32), there is more involved. The precise relationship is

$$\boxed{\sum_{m=-\infty}^{n} x[m] \overset{\mathfrak{F}}{\longleftrightarrow} \frac{1}{1-e^{-j\omega}}X(e^{j\omega}) + \pi X(e^{j0}) \sum_{k=-\infty}^{+\infty} \delta(\omega - 2\pi k).} \tag{5.39}$$

The impulse train on the right-hand side of eq. (5.39) reflects the dc or average value that can result from summation.

Example 5.8

Let us derive the Fourier transform $X(e^{j\omega})$ of the unit step $x[n] = u[n]$ by making use of the accumulation property and the knowledge that

$$g[n] = \delta[n] \overset{\mathfrak{F}}{\longleftrightarrow} G(e^{j\omega}) = 1.$$

From Section 1.4.1 we know that the unit step is the running sum of the unit impulse. That is,

$$x[n] = \sum_{m=-\infty}^{n} g[m].$$

Taking the Fourier transform of both sides and using accumulation yields

$$\begin{aligned} X(e^{j\omega}) &= \frac{1}{(1-e^{-j\omega})}G(e^{j\omega}) + \pi G(e^{j0}) \sum_{k=-\infty}^{\infty} \delta(\omega - 2\pi k) \\ &= \frac{1}{1-e^{-j\omega}} + \pi \sum_{k=-\infty}^{\infty} \delta(\omega - 2\pi k). \end{aligned}$$

5.3.6 Time Reversal

Let $x[n]$ be a signal with spectrum $X(e^{j\omega})$, and consider the transform $Y(e^{j\omega})$ of $y[n] = x[-n]$. From eq. (5.9),

$$Y(e^{j\omega}) = \sum_{n=-\infty}^{+\infty} y[n]e^{-j\omega n} = \sum_{n=-\infty}^{+\infty} x[-n]e^{-j\omega n}. \tag{5.40}$$

Substituting $m = -n$ into eq. (5.40), we obtain

$$Y(e^{j\omega}) = \sum_{m=-\infty}^{+\infty} x[m]e^{-j(-\omega)m} = X(e^{-j\omega}). \tag{5.41}$$

That is,

$$\boxed{x[-n] \overset{\mathfrak{F}}{\longleftrightarrow} X(e^{-j\omega}).} \tag{5.42}$$

5.3.7 Time Expansion

Because of the discrete nature of the time index for discrete-time signals, the relation between time and frequency scaling in discrete time takes on a somewhat different form from its continuous-time counterpart. Specifically, in Section 4.3.5 we derived the continuous-time property

$$x(at) \overset{\mathfrak{F}}{\longleftrightarrow} \frac{1}{|a|}X\left(\frac{j\omega}{a}\right). \tag{5.43}$$

However, if we try to define the signal $x[an]$, we run into difficulties if a is not an integer. Therefore, we cannot slow down the signal by choosing $a < 1$. On the other hand, if we let a be an integer other than ± 1—for example, if we consider $x[2n]$—we do not merely speed up the original signal. That is, since n can take on only integer values, the signal $x[2n]$ consists of the even samples of $x[n]$ alone.

There is a result that does closely parallel eq. (5.43), however. Let k be a positive integer, and define the signal

$$x_{(k)}[n] = \begin{cases} x[n/k], & \text{if } n \text{ is a multiple of } k \\ 0, & \text{if } n \text{ is not a multiple of } k. \end{cases} \tag{5.44}$$

As illustrated in Figure 5.13 for $k = 3$, $x_{(k)}[n]$ is obtained from $x[n]$ by placing $k-1$ zeros between successive values of the original signal. Intuitively, we can think of $x_{(k)}[n]$ as a slowed-down version of $x[n]$. Since $x_{(k)}[n]$ equals 0 unless n is a multiple of k, i.e., unless $n = rk$, we see that the Fourier transform of $x_{(k)}[n]$ is given by

$$X_{(k)}(e^{j\omega}) = \sum_{n=-\infty}^{+\infty} x_{(k)}[n]e^{-j\omega n} = \sum_{r=-\infty}^{+\infty} x_{(k)}[rk]e^{-j\omega rk}.$$

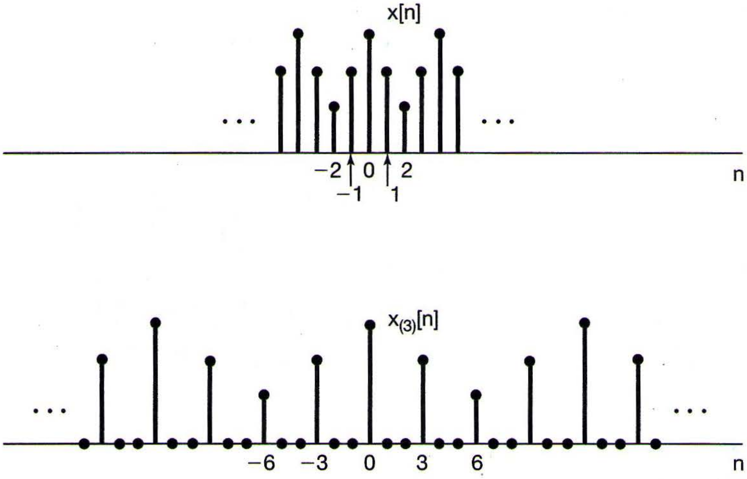


Figure 5.13 The signal $x_{(3)}[n]$ obtained from $x[n]$ by inserting two zeros between successive values of the original signal.

Furthermore, since $x_{(k)}[rk] = x[r]$, we find that

$$X_{(k)}(e^{j\omega}) = \sum_{r=-\infty}^{+\infty} x[r]e^{-j(k\omega)r} = X(e^{jk\omega}).$$

That is,

$$\boxed{x_{(k)}[n] \overset{\mathcal{F}}{\longleftrightarrow} X(e^{jk\omega}).} \tag{5.45}$$

Note that as the signal is spread out and slowed down in time by taking $k > 1$, its Fourier transform is compressed. For example, since $X(e^{j\omega})$ is periodic with period 2π, $X(e^{jk\omega})$ is periodic with period $2\pi/k$. This property is illustrated in Figure 5.14 for a rectangular pulse.

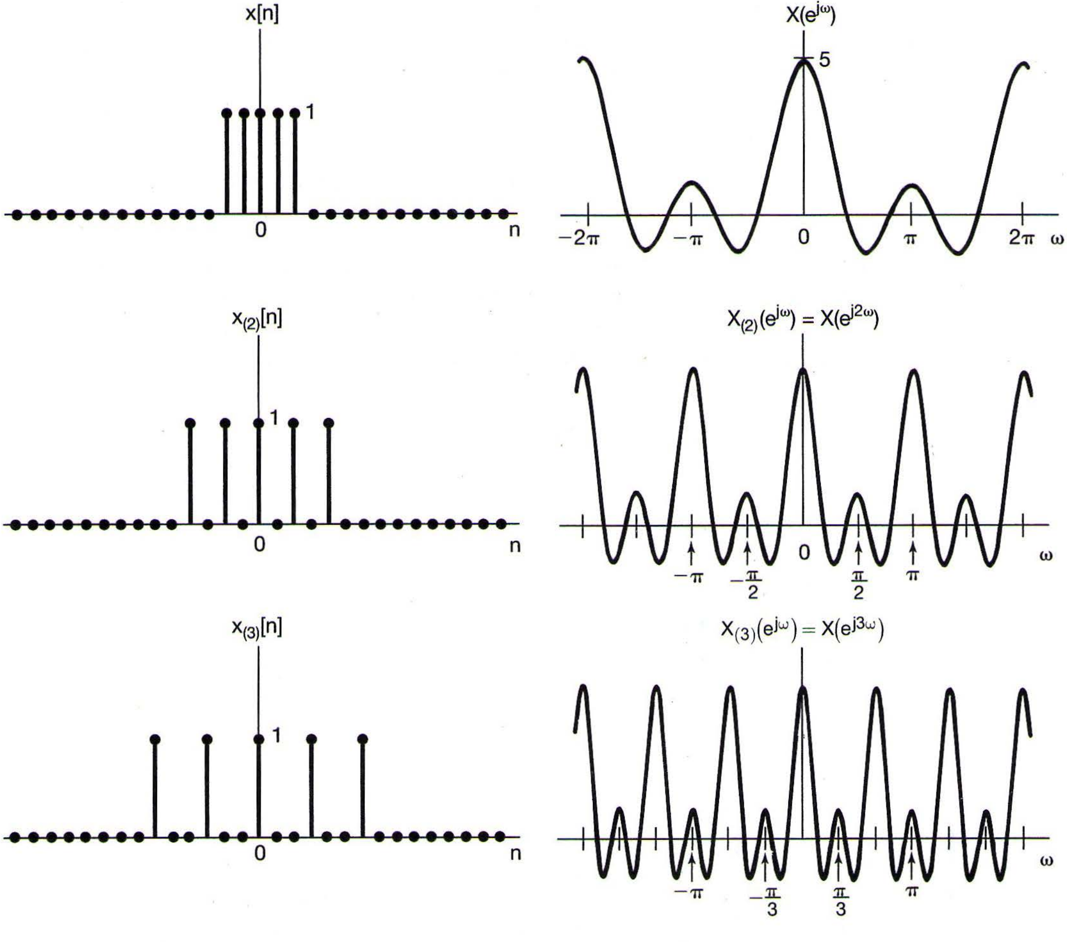


Figure 5.14 Inverse relationship between the time and frequency domains: As k increases, $x_{(k)}[n]$ spreads out while its transform is compressed.

Example 5.9

As an illustration of the usefulness of the time-expansion property in determining Fourier transforms, let us consider the sequence $x[n]$ displayed in Figure 5.15(a). This sequence can be related to the simpler sequence $y[n]$ depicted in Figure 5.15(b). In particular

$$x[n] = y_{(2)}[n] + 2y_{(2)}[n-1],$$

where

$$y_{(2)}[n] = \begin{cases} y[n/2], & \text{if } n \text{ is even} \\ 0, & \text{if } n \text{ is odd} \end{cases}$$

and $y_{(2)}[n-1]$ represents $y_{(2)}[n]$ shifted one unit to the right. The signals $y_{(2)}[n]$ and $2y_{(2)}[n-1]$ are depicted in Figures 5.15(c) and (d), respectively.

Next, note that $y[n] = g[n-2]$, where $g[n]$ is a rectangular pulse as considered in Example 5.3 (with $N_1 = 2$) and as depicted in Figure 5.6(a). Consequently, from Example 5.3 and the time-shifting property, we see that

$$Y(e^{j\omega}) = e^{-j2\omega}\frac{\sin(5\omega/2)}{\sin(\omega/2)}.$$

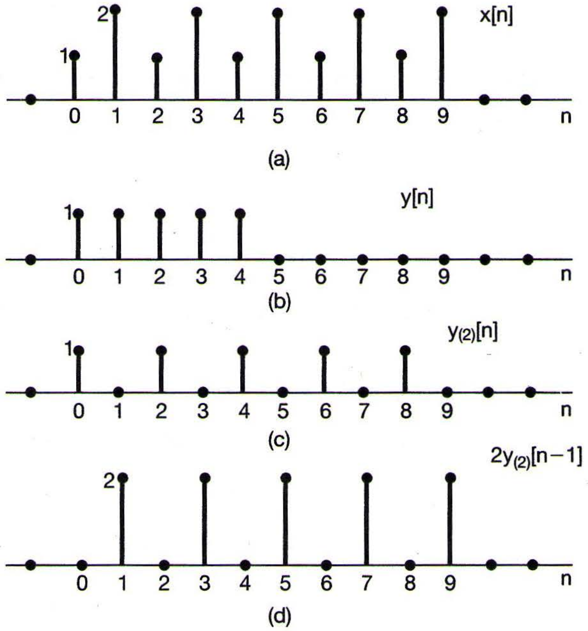


Figure 5.15 (a) The signal $x[n]$ in Example 5.9; (b) the signal $y[n]$; (c) the signal $y_{(2)}[n]$ obtained by inserting one zero between successive values of $y[n]$; and (d) the signal $2y_{(2)}[n-1]$.

Using the time-expansion property, we then obtain

$$y_{(2)}[n] \overset{\mathfrak{F}}{\longleftrightarrow} e^{-j4\omega}\frac{\sin(5\omega)}{\sin(\omega)},$$

and using the linearity and time-shifting properties, we get

$$2y_{(2)}[n-1] \overset{\mathfrak{F}}{\longleftrightarrow} 2e^{-j5\omega}\frac{\sin(5\omega)}{\sin(\omega)}.$$

Combining these two results, we have

$$X(e^{j\omega}) = e^{-j4\omega}(1+2e^{-j\omega})\left(\frac{\sin(5\omega)}{\sin(\omega)}\right).$$

5.3.8 Differentiation in Frequency

Again, let

$$x[n] \overset{\mathfrak{F}}{\longleftrightarrow} X(e^{j\omega}).$$

If we use the definition of $X(e^{j\omega})$ in the analysis equation (5.9) and differentiate both sides, we obtain

$$\frac{dX(e^{j\omega})}{d\omega} = \sum_{n=-\infty}^{+\infty} -jnx[n]e^{-j\omega n}.$$

The right-hand side of this equation is the Fourier transform of $-jnx[n]$. Therefore, multiplying both sides by j, we see that

$$\boxed{nx[n] \overset{\mathfrak{F}}{\longleftrightarrow} j\frac{dX(e^{j\omega})}{d\omega}.} \tag{5.46}$$

The usefulness of this property will be illustrated in Example 5.13 in Section 5.4.

5.3.9 Parseval's Relation

If $x[n]$ and $X(e^{j\omega})$ are a Fourier transform pair, then

$$\boxed{\sum_{n=-\infty}^{+\infty} |x[n]|^2 = \frac{1}{2\pi}\int_{2\pi} |X(e^{j\omega})|^2 d\omega.} \tag{5.47}$$

We note that this is similar to eq. (4.43), and the derivation proceeds in a similar manner. The quantity on the left-hand side of eq. (5.47) is the total energy in the signal $x[n]$, and

Parseval's relation states that this energy can also be determined by integrating the energy per unit frequency, $|X(e^{j\omega})|^2/2\pi$, over a full 2π interval of distinct discrete-time frequencies. In analogy with the continuous-time case, $|X(e^{j\omega})|^2$ is referred to as the *energy-density spectrum* of the signal $x[n]$. Note also that eq. (5.47) is the counterpart for aperiodic signals of Parseval's relation, eq. (3.110), for periodic signals, which equates the average power in a periodic signal with the sum of the average powers of its individual harmonic components.

Given the Fourier transform of a sequence, it is possible to use Fourier transform properties to determine whether a particular sequence has a number of different properties. To illustrate this idea, we present the following example.

Example 5.10

Consider the sequence $x[n]$ whose Fourier transform $X(e^{j\omega})$ is depicted for $-\pi \leq \omega \leq \pi$ in Figure 5.16. We wish to determine whether or not, in the time domain, $x[n]$ is periodic, real, even, and/or of finite energy.

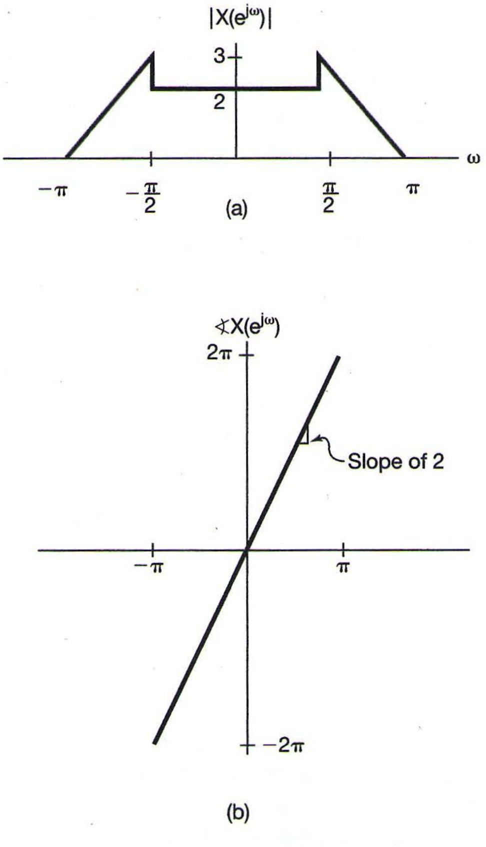


Figure 5.16 Magnitude and phase of the Fourier transform for Example 5.10.

Accordingly, we note first that periodicity in the time domain implies that the Fourier transform is zero, except possibly for impulses located at various integer multiples of the fundamental frequency. This is not true for $X(e^{j\omega})$. We conclude, then, that $x[n]$ is *not* periodic.

Next, from the symmetry properties for Fourier transforms, we know that a real-valued sequence must have a Fourier transform of even magnitude and a phase function that is odd. This is true for the given $|X(e^{j\omega})|$ and $\sphericalangle X(e^{j\omega})$. We thus conclude that $x[n]$ is real.

Third, if $x[n]$ is an even function, then, by the symmetry properties for real signals, $X(e^{j\omega})$ must be real and even. However, since $X(e^{j\omega}) = |X(e^{j\omega})|e^{-j2\omega}$, $X(e^{j\omega})$ is not a real-valued function. Consequently, $x[n]$ is not even.

Finally, to test for the finite-energy property, we may use Parseval's relation,

$$\sum_{n=-\infty}^{\infty} |x[n]|^2 = \frac{1}{2\pi}\int_{2\pi} |X(e^{j\omega})|^2 d\omega.$$

It is clear from Figure 5.16 that integrating $|X(e^{j\omega})|^2$ from $-\pi$ to π will yield a finite quantity. We conclude that $x[n]$ has finite energy.

In the next few sections, we consider several additional properties. The first two of these are the convolution and multiplication properties, similar to those discussed in Sections 4.4 and 4.5. The third is the property of duality, which is examined in Section 5.7, where we consider not only duality in the discrete-time domain, but also the duality that exists *between* the continuous-time and discrete-time domains.

5.4 THE CONVOLUTION PROPERTY

In Section 4.4, we discussed the importance of the continuous-time Fourier transform with regard to its effect on the operation of convolution and its use in dealing with continuous-time LTI systems. An identical relation applies in discrete time, and this is one of the principal reasons that the discrete-time Fourier transform is of such great value in representing and analyzing discrete-time LTI systems. Specifically, if $x[n]$, $h[n]$, and $y[n]$ are the input, impulse response, and output, respectively, of an LTI system, so that

$$y[n] = x[n] * h[n],$$

then

$$\boxed{Y(e^{j\omega}) = X(e^{j\omega})H(e^{j\omega}),} \tag{5.48}$$

where $X(e^{j\omega})$, $H(e^{j\omega})$, and $Y(e^{j\omega})$ are the Fourier transforms of $x[n]$, $h[n]$, and $y[n]$, respectively. Furthermore, comparing eqs. (3.122) and (5.9), we see that the frequency response of a discrete-time LTI system, as first defined in Section 3.8, is the Fourier transform of the impulse response of the system.

The derivation of eq. (5.48) exactly parallels that carried out in Section 4.4. In particular, as in continuous time, the Fourier synthesis equation (5.8) for $x[n]$ can be interpreted as a decomposition of $x[n]$ into a linear combination of complex exponentials with infinitesimal amplitudes proportional to $X(e^{j\omega})$. Each of these exponentials is an eigenfunction of the system. In Chapter 3, we used this fact to show that the Fourier series coefficients of the response of an LTI system to a periodic input are simply the Fourier coefficients of the input multiplied by the system's frequency response evaluated at the corresponding harmonic frequencies. The convolution property (5.48) represents the extension of this result to aperiodic inputs and outputs by using the Fourier transform rather than the Fourier series.

As in continuous time, eq. (5.48) maps the convolution of two signals to the simple algebraic operation of multiplying their Fourier transforms, a fact that both facilitates the analysis of signals and systems and adds significantly to our understanding of the way in which an LTI system responds to the input signals that are applied to it. In particular, from eq. (5.48), we see that the frequency response $H(e^{j\omega})$ captures the change in complex amplitude of the Fourier transform of the input at each frequency ω. Thus, in frequency-selective filtering, for example, we want $H(e^{j\omega}) \approx 1$ over the range of frequencies corresponding to the desired passband and $H(e^{j\omega}) \approx 0$ over the band of frequencies to be eliminated or significantly attenuated.

5.4.1 Examples

To illustrate the convolution property, along with a number of other properties, we consider several examples in this section.

Example 5.11

Consider an LTI system with impulse response

$$h[n] = \delta[n - n_0].$$

The frequency response is

$$H(e^{j\omega}) = \sum_{n=-\infty}^{+\infty} \delta[n - n_0]e^{-j\omega n} = e^{-j\omega n_0}.$$

Thus, for any input $x[n]$ with Fourier transform $X(e^{j\omega})$, the Fourier transform of the output is

$$Y(e^{j\omega}) = e^{-j\omega n_0} X(e^{j\omega}). \tag{5.49}$$

We note that, for this example, $y[n] = x[n - n_0]$ and eq. (5.49) is consistent with the time-shifting property. Note also that the frequency response $H(e^{j\omega}) = e^{-j\omega n_0}$ of a pure time shift has unity magnitude at all frequencies and a phase characteristic $-\omega n_0$ that is linear with frequency.

Example 5.12

Consider the discrete-time ideal lowpass filter introduced in Section 3.9.2. This system has the frequency response $H(e^{j\omega})$ illustrated in Figure 5.17(a). Since the impulse response and frequency response of an LTI system are a Fourier transform pair, we can determine the impulse response of the ideal lowpass filter from the frequency response using the Fourier transform synthesis equation (5.8). In particular, using $-\pi \leq \omega \leq \pi$ as the interval of integration in that equation, we see from Figure 5.17(a) that

$$\begin{aligned} h[n] &= \frac{1}{2\pi}\int_{-\pi}^{\pi} H(e^{j\omega})e^{j\omega n}d\omega = \frac{1}{2\pi}\int_{-\omega_c}^{\omega_c} e^{j\omega n}d\omega \\ &= \frac{\sin \omega_c n}{\pi n}, \end{aligned} \tag{5.50}$$

which is shown in Figure 5.17(b).

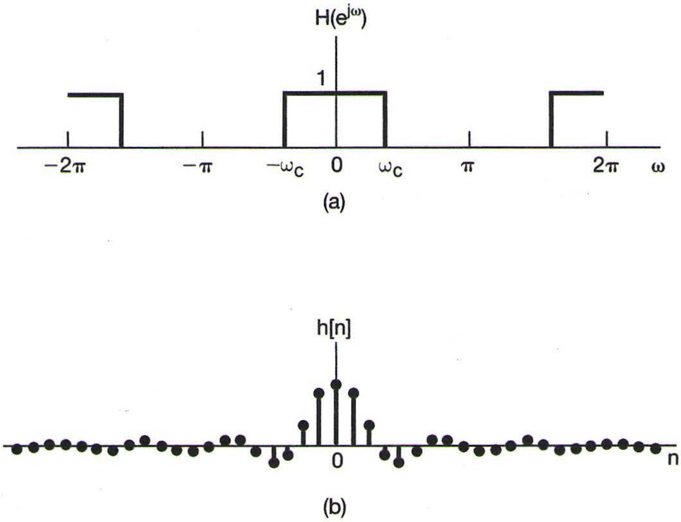


Figure 5.17 (a) Frequency response of a discrete-time ideal lowpass filter; (b) impulse response of the ideal lowpass filter.

In Figure 5.17, we come across many of the same issues that surfaced with the continuous-time ideal lowpass filter in Example 4.18. First, since $h[n]$ is not zero for $n < 0$, the ideal lowpass filter is not causal. Second, even if causality is not an important issue, there are other reasons, including ease of implementation and preferable time domain characteristics, that nonideal filters are generally used to perform frequency-selective filtering. In particular, the impulse response of the ideal lowpass filter in Figure 5.17(b) is oscillatory, a characteristic that is undesirable in some applications. In such cases, a tradeoff between frequency-domain objectives such as frequency selectivity and time-domain properties such as nonoscillatory behavior must be made. In Chapter 6, we will discuss these and related ideas in more detail.

As the following example illustrates, the convolution property can also be of value in facilitating the calculation of convolution sums.

Example 5.13

Consider an LTI system with impulse response

$$h[n] = \alpha^n u[n],$$

with $|\alpha| < 1$, and suppose that the input to this system is

$$x[n] = \beta^n u[n],$$

with $|\beta| < 1$. Evaluating the Fourier transforms of $h[n]$ and $x[n]$, we have

$$H(e^{j\omega}) = \frac{1}{1 - \alpha e^{-j\omega}} \tag{5.51}$$

and

$$X(e^{j\omega}) = \frac{1}{1 - \beta e^{-j\omega}}, \tag{5.52}$$

so that

$$Y(e^{j\omega}) = H(e^{j\omega})X(e^{j\omega}) = \frac{1}{(1 - \alpha e^{-j\omega})(1 - \beta e^{-j\omega})}. \tag{5.53}$$

As with Example 4.19 , determining the inverse transform of $Y(e^{j\omega})$ is most easily done by expanding $Y(e^{j\omega})$ by the method of partial fractions. Specifically, $Y(e^{j\omega})$ is a ratio of polynomials in powers of $e^{-j\omega}$, and we would like to express this as a sum of simpler terms of this type so that we can find the inverse transform of each term by inspection (together, perhaps, with the use of the frequency differentiation property of Section 5.3.8). The general algebraic procedure for rational transforms is described in the appendix. For this example, if $\alpha \neq \beta$, the partial fraction expansion of $Y(e^{j\omega})$ is of the form

$$Y(e^{j\omega}) = \frac{A}{1 - \alpha e^{-j\omega}} + \frac{B}{1 - \beta e^{-j\omega}}. \tag{5.54}$$

Equating the right-hand sides of eqs (5.53) and (5.54), we find that

$$A = \frac{\alpha}{\alpha - \beta}, \quad B = -\frac{\beta}{\alpha - \beta}.$$

Therefore, from Example 5.1 and the linearity property, we can obtain the inverse transform of eq. (5.54) by inspection:

$$\begin{aligned} y[n] &= \frac{\alpha}{\alpha - \beta}\alpha^n u[n] - \frac{\beta}{\alpha - \beta}\beta^n u[n] \\ &= \frac{1}{\alpha - \beta}[\alpha^{n+1}u[n] - \beta^{n+1}u[n]]. \end{aligned} \tag{5.55}$$

For $\alpha = \beta$, the partial-fraction expansion in eq. (5.54) is not valid. However, in this case,

$$Y(e^{j\omega}) = \left(\frac{1}{1 - \alpha e^{-j\omega}}\right)^2,$$

which can be expressed as

$$Y(e^{j\omega}) = \frac{j}{\alpha}e^{j\omega}\frac{d}{d\omega}\left(\frac{1}{1-\alpha e^{-j\omega}}\right). \tag{5.56}$$

As in Example 4.19, we can use the frequency differentiation property, eq. (5.46), together with the Fourier transform pair

$$\alpha^n u[n] \overset{\mathfrak{F}}{\longleftrightarrow} \frac{1}{1-\alpha e^{-j\omega}},$$

to conclude that

$$n\alpha^n u[n] \overset{\mathfrak{F}}{\longleftrightarrow} j\frac{d}{d\omega}\left(\frac{1}{1-\alpha e^{-j\omega}}\right).$$

To account for the factor $e^{j\omega}$, we use the time-shifting property to obtain

$$(n+1)\alpha^{n+1}u[n+1] \overset{\mathfrak{F}}{\longleftrightarrow} je^{j\omega}\frac{d}{d\omega}\left(\frac{1}{1-\alpha e^{-j\omega}}\right),$$

and finally, accounting for the factor $1/\alpha$, in eq. (5.56), we obtain

$$y[n] = (n+1)\alpha^n u[n+1]. \tag{5.57}$$

It is worth noting that, although the right-hand side is multiplied by a step that begins at $n = -1$, the sequence $(n+1)\alpha^n u[n+1]$ is still zero prior to $n = 0$, since the factor $n+1$ is zero at $n = -1$. Thus, we can alternatively express $y[n]$ as

$$y[n] = (n+1)\alpha^n u[n]. \tag{5.58}$$

As illustrated in the next example, the convolution property, along with other Fourier transform properties, is often useful in analyzing system interconnections.

Example 5.14

Consider the system shown in Figure 5.18(a) with input $x[n]$ and output $y[n]$. The LTI systems with frequency response $H_{lp}(e^{j\omega})$ are ideal lowpass filters with cutoff frequency $\pi/4$ and unity gain in the passband.

Let us first consider the top path in Figure 5.18(a). The Fourier transform of the signal $w_1[n]$ can be obtained by noting that $(-1)^n = e^{j\pi n}$ so that $w_1[n] = e^{j\pi n}x[n]$. Using the frequency-shifting property, we then obtain

$$W_1(e^{j\omega}) = X(e^{j(\omega-\pi)}).$$

The convolution property yields

$$W_2(e^{j\omega}) = H_{lp}(e^{j\omega})X(e^{j(\omega-\pi)}).$$

Since $w_3[n] = e^{j\pi n}w_2[n]$, we can again apply the frequency-shifting property to obtain

$$\begin{aligned} W_3(e^{j\omega}) &= W_2(e^{j(\omega-\pi)}) \\ &= H_{lp}(e^{j(\omega-\pi)})X(e^{j(\omega-2\pi)}). \end{aligned}$$

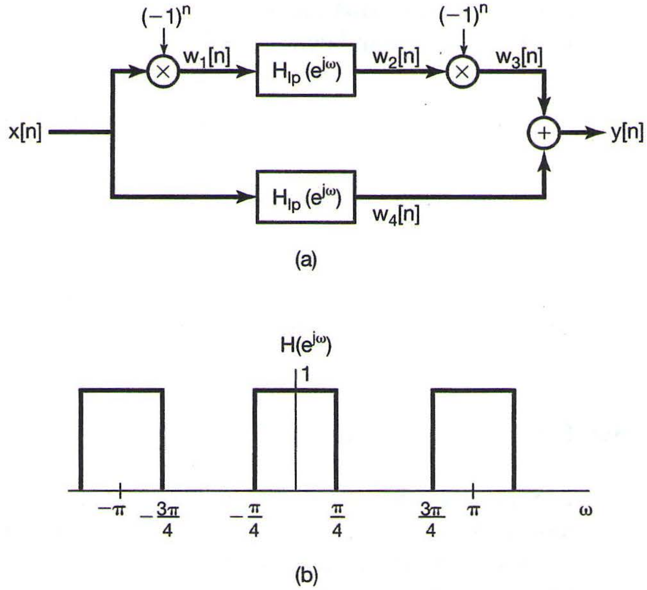


Figure 5.18 (a) System interconnection for Example 5.14; (b) the overall frequency response for this system.

Since discrete-time Fourier transforms are always periodic with period 2π,

$$W_3(e^{j\omega}) = H_{lp}(e^{j(\omega-\pi)})X(e^{j\omega}).$$

Applying the convolution property to the lower path, we get

$$W_4(e^{j\omega}) = H_{lp}(e^{j\omega})X(e^{j\omega}).$$

From the linearity property of the Fourier transform, we obtain

$$\begin{aligned} Y(e^{j\omega}) &= W_3(e^{j\omega}) + W_4(e^{j\omega}) \\ &= [H_{lp}(e^{j(\omega-\pi)}) + H_{lp}(e^{j\omega})]X(e^{j\omega}). \end{aligned}$$

Consequently, the overall system in Figure 5.18(a) has the frequency response

$$H(e^{j\omega}) = [H_{lp}(e^{j(\omega-\pi)}) + H_{lp}(e^{j\omega})]$$

which is shown in Figure 5.18(b).

As we saw in Example 5.7, $H_{lp}(e^{j(\omega-\pi)})$ is the frequency response of an ideal highpass filter. Thus, the overall system passes both low and high frequencies and stops frequencies between these two passbands. That is, the filter has what is often referred to as an *ideal bandstop characteristic,* where the stopband is the region $\pi/4 < |\omega| < 3\pi/4$.

It is important to note that, as in continuous time, not every discrete-time LTI system has a frequency response. For example, the LTI system with impulse response $h[n] = 2^n u[n]$ does not have a finite response to sinusoidal inputs, which is reflected in the fact that the Fourier transform analysis equation for $h[n]$ diverges. However, if an LTI system is stable, then, from Section 2.3.7, its impulse response is absolutely summable; that is,

$$\sum_{n=-\infty}^{+\infty} |h[n]| < \infty. \tag{5.59}$$

Therefore, the frequency response always converges for stable systems. In using Fourier methods, we will be restricting ourselves to systems with impulse responses that have well-defined Fourier transforms. In Chapter 10, we will introduce an extension of the Fourier transform referred to as the z-transform that will allow us to use transform techniques for LTI systems for which the frequency response does not converge.

5.5 THE MULTIPLICATION PROPERTY

In Section 4.5, we introduced the multiplication property for continuous-time signals and indicated some of its applications through several examples. An analogous property exists for discrete-time signals and plays a similar role in applications. In this section, we derive this result directly and give an example of its use. In Chapters 7 and 8, we will use the multiplication property in the context of our discussions of sampling and communications.

Consider $y[n]$ equal to the product of $x_1[n]$ and $x_2[n]$, with $Y(e^{j\omega})$, $X_1(e^{j\omega})$, and $X_2(e^{j\omega})$ denoting the corresponding Fourier transforms. Then

$$Y(e^{j\omega}) = \sum_{n=-\infty}^{+\infty} y[n]e^{-j\omega n} = \sum_{n=-\infty}^{+\infty} x_1[n]x_2[n]e^{-j\omega n},$$

or since

$$x_1[n] = \frac{1}{2\pi}\int_{2\pi} X_1(e^{j\theta})e^{j\theta n}d\theta, \tag{5.60}$$

it follows that

$$Y(e^{j\omega}) = \sum_{n=-\infty}^{+\infty} x_2[n]\left\{\frac{1}{2\pi}\int_{2\pi} X_1(e^{j\theta})e^{j\theta n}d\theta\right\} e^{-j\omega n}. \tag{5.61}$$

Interchanging the order of summation and integration, we obtain

$$Y(e^{j\omega}) = \frac{1}{2\pi}\int_{2\pi} X_1(e^{j\theta})\left[\sum_{n=-\infty}^{+\infty} x_2[n]e^{-j(\omega-\theta)n}\right]d\theta. \tag{5.62}$$

The bracketed summation is $X_2(e^{j(\omega-\theta)})$, and consequently, eq. (5.62) becomes

$$\boxed{Y(e^{j\omega}) = \frac{1}{2\pi}\int_{2\pi} X_1(e^{j\theta})X_2(e^{j(\omega-\theta)})d\theta.} \tag{5.63}$$

Equation (5.63) corresponds to a *periodic* convolution of $X_1(e^{j\omega})$ and $X_2(e^{j\omega})$, and the integral in this equation can be evaluated over any interval of length 2π. The usual form of convolution (in which the integral ranges from $-\infty$ to $+\infty$) is often referred to as aperiodic convolution to distinguish it from periodic convolution. The mechanics of periodic convolution are most easily illustrated through an example.

Example 5.15

Consider the problem of finding the Fourier transform $X(e^{j\omega})$ of a signal $x[n]$ which is the product of two other signals; that is,

$$x[n] = x_1[n]x_2[n],$$

where

$$x_1[n] = \frac{\sin(3\pi n/4)}{\pi n}$$

and

$$x_2[n] = \frac{\sin(\pi n/2)}{\pi n}.$$

From the multiplication property given in eq. (5.63), we know that $X(e^{j\omega})$ is the periodic convolution of $X_1(e^{j\omega})$ and $X_2(e^{j\omega})$, where the integral in eq. (5.63) can be taken over any interval of length 2π. Choosing the interval $-\pi < \theta \leq \pi$, we obtain

$$X(e^{j\omega}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} X_1(e^{j\theta})X_2(e^{j(\omega-\theta)})d\theta. \tag{5.64}$$

Equation (5.64) resembles aperiodic convolution, except for the fact that the integration is limited to the interval $-\pi < \theta \leq \pi$. However, we can convert the equation into an ordinary convolution by defining

$$\hat{X}_1(e^{j\omega}) = \begin{cases} X_1(e^{j\omega}) & \text{for } -\pi < \omega \leq \pi \\ 0 & \text{otherwise} \end{cases}.$$

Then, replacing $X_1(e^{j\theta})$ in eq. (5.64) by $\hat{X}_1(e^{j\theta})$, and using the fact that $\hat{X}_1(e^{j\theta})$ is zero for $|\theta| > \pi$, we see that

$$\begin{aligned} X(e^{j\omega}) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \hat{X}_1(e^{j\theta})X_2(e^{j(\omega-\theta)})d\theta \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{X}_1(e^{j\theta})X_2(e^{j(\omega-\theta)})d\theta. \end{aligned}$$

Thus, $X(e^{j\omega})$ is $1/2\pi$ times the aperiodic convolution of the rectangular pulse $\hat{X}_1(e^{j\omega})$ and the periodic square wave $X_2(e^{j\omega})$, both of which are shown in Figure 5.19. The result of this convolution is the Fourier transform $X(e^{j\omega})$ shown in Figure 5.20.

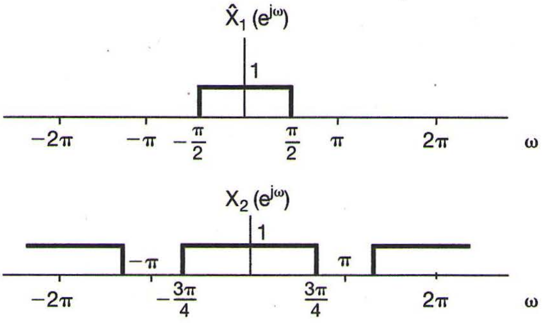


Figure 5.19 $\hat{X}_1(e^{j\omega})$ representing one period of $X_1(e^{j\omega})$, and $X_2(e^{j\omega})$. The linear convolution of $\hat{X}_1(e^{j\omega})$ and $X_2(e^{j\omega})$ corresponds to the periodic convolution of $X_1(e^{j\omega})$ and $X_2(e^{j\omega})$.

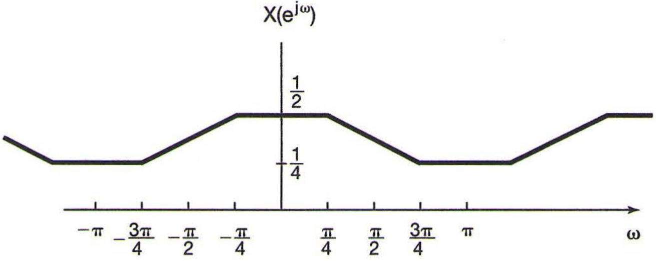


Figure 5.20 Result of the periodic convolution in Example 5.15.

5.6 TABLES OF FOURIER TRANSFORM PROPERTIES AND BASIC FOURIER TRANSFORM PAIRS

In Table 5.1, we summarize a number of important properties of the discrete-time Fourier transform and indicate the section of the text in which each is discussed. In Table 5.2, we summarize some of the basic and most important discrete-time Fourier transform pairs. Many of these have been derived in examples in the chapter.

5.7 DUALITY

In considering the continuous-time Fourier transform, we observed a symmetry or duality between the analysis equation (4.9) and the synthesis equation (4.8). No corresponding duality exists between the analysis equation (5.9) and the synthesis equation (5.8) for the discrete-time Fourier transform. However, there *is* a duality in the discrete-time Fourier *series* equations (3.94) and (3.95), which we develop in Section 5.7.1. In addition, there is

TABLE 5.1 PROPERTIES OF THE DISCRETE-TIME FOURIER TRANSFORM

Section	Property	Aperiodic Signal	Fourier Transform
		$x[n]$ $y[n]$	$\left.\begin{array}{l}X(e^{j\omega})\\Y(e^{j\omega})\end{array}\right\}$ periodic with period 2π
5.3.2	Linearity	$ax[n] + by[n]$	$aX(e^{j\omega}) + bY(e^{j\omega})$
5.3.3	Time Shifting	$x[n - n_0]$	$e^{-j\omega n_0}X(e^{j\omega})$
5.3.3	Frequency Shifting	$e^{j\omega_0 n}x[n]$	$X(e^{j(\omega-\omega_0)})$
5.3.4	Conjugation	$x^*[n]$	$X^*(e^{-j\omega})$
5.3.6	Time Reversal	$x[-n]$	$X(e^{-j\omega})$
5.3.7	Time Expansion	$x_{(k)}[n] = \begin{cases} x[n/k], & \text{if } n = \text{multiple of } k \\ 0, & \text{if } n \neq \text{multiple of } k \end{cases}$	$X(e^{jk\omega})$
5.4	Convolution	$x[n] * y[n]$	$X(e^{j\omega})Y(e^{j\omega})$
5.5	Multiplication	$x[n]y[n]$	$\frac{1}{2\pi}\int_{2\pi} X(e^{j\theta})Y(e^{j(\omega-\theta)})d\theta$
5.3.5	Differencing in Time	$x[n] - x[n-1]$	$(1 - e^{-j\omega})X(e^{j\omega})$
5.3.5	Accumulation	$\sum_{k=-\infty}^{n} x[k]$	$\frac{1}{1 - e^{-j\omega}}X(e^{j\omega})$ $+\pi X(e^{j0}) \sum_{k=-\infty}^{+\infty} \delta(\omega - 2\pi k)$
5.3.8	Differentiation in Frequency	$nx[n]$	$j\frac{dX(e^{j\omega})}{d\omega}$
5.3.4	Conjugate Symmetry for Real Signals	$x[n]$ real	$\begin{cases} X(e^{j\omega}) = X^*(e^{-j\omega}) \\ \mathcal{R}e\{X(e^{j\omega})\} = \mathcal{R}e\{X(e^{-j\omega})\} \\ \mathcal{I}m\{X(e^{j\omega})\} = -\mathcal{I}m\{X(e^{-j\omega})\} \\ \lvert X(e^{j\omega})\rvert = \lvert X(e^{-j\omega})\rvert \\ \sphericalangle X(e^{j\omega}) = -\sphericalangle X(e^{-j\omega}) \end{cases}$
5.3.4	Symmetry for Real, Even Signals	$x[n]$ real an even	$X(e^{j\omega})$ real and even
5.3.4	Symmetry for Real, Odd Signals	$x[n]$ real and odd	$X(e^{j\omega})$ purely imaginary and odd
5.3.4	Even-odd Decomposition of Real Signals	$x_e[n] = \mathcal{E}v\{x[n]\}$ $[x[n]$ real$]$ $x_o[n] = \mathcal{O}d\{x[n]\}$ $[x[n]$ real$]$	$\mathcal{R}e\{X(e^{j\omega})\}$ $j\mathcal{I}m\{X(e^{j\omega})\}$
5.3.9	Parseval's Relation for Aperiodic Signals	$\sum_{n=-\infty}^{+\infty} \lvert x[n]\rvert^2 = \frac{1}{2\pi}\int_{2\pi} \lvert X(e^{j\omega})\rvert^2 d\omega$	

a duality relationship between the discrete-time Fourier transform and the continuous-time Fourier series. This relation is discussed in Section 5.7.2.

5.7.1 Duality in the Discrete-Time Fourier Series

Since the Fourier series coefficients a_k of a periodic signal $x[n]$ are themselves a periodic sequence, we can expand the sequence a_k in a Fourier series. The duality property for discrete-time Fourier series implies that the Fourier series coefficients for the periodic sequence a_k are the values of $(1/N)x[-n]$ (i.e., are proportional to the values of the original

TABLE 5.2 BASIC DISCRETE-TIME FOURIER TRANSFORM PAIRS

Signal	**Fourier Transform**	**Fourier Series Coefficients (if periodic)**
$\sum_{k=\langle N\rangle} a_k e^{jk(2\pi/N)n}$	$2\pi \sum_{k=-\infty}^{+\infty} a_k \delta\left(\omega - \frac{2\pi k}{N}\right)$	a_k
$e^{j\omega_0 n}$	$2\pi \sum_{l=-\infty}^{+\infty} \delta(\omega - \omega_0 - 2\pi l)$	(a) $\omega_0 = \frac{2\pi m}{N}$ $a_k = \begin{cases} 1, & k = m, m \pm N, m \pm 2N, \dots \\ 0, & \text{otherwise} \end{cases}$ (b) $\frac{\omega_0}{2\pi}$ irrational $\Rightarrow$ The signal is aperiodic
$\cos \omega_0 n$	$\pi \sum_{l=-\infty}^{+\infty} \{\delta(\omega - \omega_0 - 2\pi l) + \delta(\omega + \omega_0 - 2\pi l)\}$	(a) $\omega_0 = \frac{2\pi m}{N}$ $a_k = \begin{cases} \frac{1}{2}, & k = \pm m, \pm m \pm N, \pm m \pm 2N, \dots \\ 0, & \text{otherwise} \end{cases}$ (b) $\frac{\omega_0}{2\pi}$ irrational $\Rightarrow$ The signal is aperiodic
$\sin \omega_0 n$	$\frac{\pi}{j} \sum_{l=-\infty}^{+\infty} \{\delta(\omega - \omega_0 - 2\pi l) - \delta(\omega + \omega_0 - 2\pi l)\}$	(a) $\omega_0 = \frac{2\pi r}{N}$ $a_k = \begin{cases} \frac{1}{2j}, & k = r, r \pm N, r \pm 2N, \dots \\ -\frac{1}{2j}, & k = -r, -r \pm N, -r \pm 2N, \dots \\ 0, & \text{otherwise} \end{cases}$ (b) $\frac{\omega_0}{2\pi}$ irrational $\Rightarrow$ The signal is aperiodic
$x[n] = 1$	$2\pi \sum_{l=-\infty}^{+\infty} \delta(\omega - 2\pi l)$	$a_k = \begin{cases} 1, & k = 0, \pm N, \pm 2N, \dots \\ 0, & \text{otherwise} \end{cases}$
Periodic square wave $x[n] = \begin{cases} 1, & \lvert n\rvert \le N_1 \\ 0, & N_1 < \lvert n\rvert \le N/2 \end{cases}$ and $x[n+N] = x[n]$	$2\pi \sum_{k=-\infty}^{+\infty} a_k \delta\left(\omega - \frac{2\pi k}{N}\right)$	$a_k = \frac{\sin[(2\pi k/N)(N_1 + \frac{1}{2})]}{N \sin[2\pi k/2N]}, \ k \ne 0, \pm N, \pm 2N, \dots$ $a_k = \frac{2N_1 + 1}{N}, \ k = 0, \pm N, \pm 2N, \dots$
$\sum_{k=-\infty}^{+\infty} \delta[n - kN]$	$\frac{2\pi}{N} \sum_{k=-\infty}^{+\infty} \delta\left(\omega - \frac{2\pi k}{N}\right)$	$a_k = \frac{1}{N}$ for all k
$a^n u[n], \quad \lvert a\rvert < 1$	$\frac{1}{1 - ae^{-j\omega}}$	—
$x[n] = \begin{cases} 1, & \lvert n\rvert \le N_1 \\ 0, & \lvert n\rvert > N_1 \end{cases}$	$\frac{\sin[\omega(N_1 + \frac{1}{2})]}{\sin(\omega/2)}$	—
$\frac{\sin Wn}{\pi n} = \frac{W}{\pi} \operatorname{sinc}\left(\frac{Wn}{\pi}\right)$ $0 < W < \pi$	$X(\omega) = \begin{cases} 1, & 0 \le \lvert\omega\rvert \le W \\ 0, & W < \lvert\omega\rvert \le \pi \end{cases}$ $X(\omega)$ periodic with period 2π	—
$\delta[n]$	1	—
$u[n]$	$\frac{1}{1 - e^{-j\omega}} + \sum_{k=-\infty}^{+\infty} \pi\delta(\omega - 2\pi k)$	—
$\delta[n - n_0]$	$e^{-j\omega n_0}$	—
$(n+1)a^n u[n], \quad \lvert a\rvert < 1$	$\frac{1}{(1 - ae^{-j\omega})^2}$	—
$\frac{(n+r-1)!}{n!(r-1)!} a^n u[n], \quad \lvert a\rvert < 1$	$\frac{1}{(1 - ae^{-j\omega})^r}$	—

signal, reversed in time). To see this in more detail, consider two periodic sequences with period N, related through the summation

$$f[m] = \frac{1}{N} \sum_{r=\langle N \rangle} g[r] e^{-jr(2\pi/N)m}. \tag{5.65}$$

If we let $m = k$ and $r = n$, eq. (5.65) becomes

$$f[k] = \frac{1}{N} \sum_{n=\langle N \rangle} g[n] e^{-jk(2\pi/N)n}.$$

Comparing this with eq. (3.95), we see that the sequence $f[k]$ corresponds to the Fourier series coefficients of the signal $g[n]$. That is, if we adopt the notation

$$x[n] \overset{\mathfrak{FS}}{\longleftrightarrow} a_k$$

introduced in Chapter 3 for a periodic discrete-time signal and its set of Fourier coefficients, the two periodic sequences related through eq. (5.65) satisfy

$$g[n] \overset{\mathfrak{FS}}{\longleftrightarrow} f[k]. \tag{5.66}$$

Alternatively, if we let $m = n$ and $r = -k$, eq. (5.65) becomes

$$f[n] = \sum_{k=\langle N \rangle} \frac{1}{N} g[-k] e^{jk(2\pi/N)n}.$$

Comparing this with eq. (3.94), we find that $(1/N)g[-k]$ corresponds to the sequence of Fourier series coefficients of $f[n]$. That is,

$$f[n] \overset{\mathfrak{FS}}{\longleftrightarrow} \frac{1}{N} g[-k]. \tag{5.67}$$

As in continuous time, this duality implies that every property of the discrete-time Fourier series has a dual. For example, referring to Table 3.2, we see that the pair of properties

$$x[n - n_0] \overset{\mathfrak{FS}}{\longleftrightarrow} a_k e^{-jk(2\pi/N)n_0} \tag{5.68}$$

and

$$e^{jm(2\pi/N)n} x[n] \overset{\mathfrak{FS}}{\longleftrightarrow} a_{k-m} \tag{5.69}$$

are dual. Similarly, from the same table, we can extract another pair of dual properties:

$$\sum_{r=\langle N \rangle} x[r] y[n - r] \overset{\mathfrak{FS}}{\longleftrightarrow} N a_k b_k \tag{5.70}$$

and

$$x[n] y[n] \overset{\mathfrak{FS}}{\longleftrightarrow} \sum_{l=\langle N \rangle} a_l b_{k-l}. \tag{5.71}$$

In addition to its consequences for the properties of discrete-time Fourier series, duality can often be useful in reducing the complexity of the calculations involved in determining Fourier series representations. This is illustrated in the following example.

Example 5.16

Consider the following periodic signal with a period of $N = 9$:

$$x[n] = \begin{cases} \frac{1}{9}\frac{\sin(5\pi n/9)}{\sin(\pi n/9)}, & n \neq \text{multiple of } 9 \\ \frac{5}{9}, & n = \text{multiple of } 9 \end{cases} \tag{5.72}$$

In Chapter 3, we found that a rectangular square wave has Fourier coefficients in a form much as in eq. (5.72). Duality, then, suggests that the coefficients for $x[n]$ must be in the form of a rectangular square wave. To see this more precisely, let $g[n]$ be a rectangular square wave with period $N = 9$ such that

$$g[n] = \begin{cases} 1, & |n| \leq 2 \\ 0, & 2 < |n| \leq 4. \end{cases}$$

The Fourier series coefficients b_k for $g[n]$ can be determined from Example 3.12 as

$$b_k = \begin{cases} \frac{1}{9}\frac{\sin(5\pi k/9)}{\sin(\pi k/9)}, & k \neq \text{multiple of } 9 \\ \frac{5}{9}, & k = \text{multiple of } 9 \end{cases}$$

The Fourier series analysis equation (3.95) for $g[n]$ can now be written as

$$b_k = \frac{1}{9}\sum_{n=-2}^{2}(1)e^{-j2\pi nk/9}.$$

Interchanging the names of the variables k and n and noting that $x[n] = b_n$, we find that

$$x[n] = \frac{1}{9}\sum_{k=-2}^{2}(1)e^{-j2\pi nk/9}.$$

Letting $k' = -k$ in the sum on the right side, we obtain

$$x[n] = \frac{1}{9}\sum_{k'=-2}^{2} e^{+j2\pi nk'/9}.$$

Finally, moving the factor 1/9 inside the summation, we see that the right side of this equation has the form of the synthesis equation (3.94) for $x[n]$. We thus conclude that the Fourier coefficients of $x[n]$ are given by

$$a_k = \begin{cases} 1/9, & |k| \leq 2 \\ 0, & 2 < |k| \leq 4, \end{cases}$$

and, of course, are periodic with period $N = 9$.

5.7.2 Duality between the Discrete-Time Fourier Transform and the Continuous-Time Fourier Series

In addition to the duality for the discrete Fourier series, there is a duality between the *discrete-time* Fourier *transform* and the *continuous-time* Fourier *series*. Specifically, let us compare the continuous-time Fourier series equations (3.38) and (3.39) with the discrete-time Fourier transform equations (5.8) and (5.9). We repeat these equations here for convenience:

[eq. (5.8)]
$$x[n] = \frac{1}{2\pi}\int_{2\pi} X(e^{j\omega})e^{j\omega n}d\omega, \tag{5.73}$$

[eq. (5.9)]
$$X(e^{j\omega}) = \sum_{n=-\infty}^{+\infty} x[n]e^{-j\omega n}, \tag{5.74}$$

[eq. (3.38)]
$$x(t) = \sum_{k=-\infty}^{+\infty} a_k e^{jk\omega_0 t}, \tag{5.75}$$

[eq. (3.39)]
$$a_k = \frac{1}{T}\int_T x(t)e^{-jk\omega_0 t}dt. \tag{5.76}$$

Note that eqs. (5.73) and (5.76) are very similar, as are eqs. (5.74) and (5.75), and in fact, we can interpret eqs. (5.73) and (5.74) as a *Fourier series* representation of the periodic frequency response $X(e^{j\omega})$. In particular, since $X(e^{j\omega})$ is a periodic function of ω with period 2π, it has a Fourier series representation as a weighted sum of harmonically related periodic exponential functions of ω, all of which have the common period of 2π. That is, $X(e^{j\omega})$ can be represented in a Fourier series as a weighted sum of the signals $e^{j\omega n}$, $n = 0, \pm 1, \pm 2, \ldots$. From eq. (5.74), we see that the nth Fourier coefficient in this expansion—i.e., the coefficient multiplying $e^{j\omega n}$—is $x[-n]$. Furthermore, since the period of $X(e^{j\omega})$ is 2π, eq. (5.73) can be interpreted as the Fourier series analysis equation for the Fourier series coefficient $x[n]$—i.e., for the coefficient multiplying $e^{-j\omega n}$ in the expression for $X(e^{j\omega})$ in eq. (5.74). The use of this duality relationship is best illustrated with an example.

Example 5.17

The duality between the discrete-time Fourier transform synthesis equation and the continuous-time Fourier series analysis equation may be exploited to determine the discrete-time Fourier transform of the sequence

$$x[n] = \frac{\sin(\pi n/2)}{\pi n}.$$

To use duality, we first must identify a continuous-time signal $g(t)$ with period $T = 2\pi$ and Fourier coefficients $a_k = x[k]$. From Example 3.5, we know that if $g(t)$ is a periodic square wave with period 2π (or, equivalently, with fundamental frequency $\omega_0 = 1$) and with

$$g(t) = \begin{cases} 1, & |t| \le T_1 \\ 0, & T_1 < |t| \le \pi \end{cases},$$

then the Fourier series coefficients of $g(t)$ are

$$a_k = \frac{\sin(kT_1)}{k\pi}$$

Consequently, if we take $T_1 = \pi/2$, we will have $a_k = x[k]$. In this case the analysis equation for $g(t)$ is

$$\frac{\sin(\pi k/2)}{\pi k} = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)e^{-jkt}dt = \frac{1}{2\pi}\int_{-\pi/2}^{\pi/2}(1)e^{-jkt}dt.$$

Renaming k as n and t as ω, we have

$$\frac{\sin(\pi n/2)}{\pi n} = \frac{1}{2\pi}\int_{-\pi/2}^{\pi/2}(1)e^{-jn\omega}d\omega. \tag{5.77}$$

Replacing n by $-n$ on both sides of eq. (5.77) and noting that the sinc function is even, we obtain

$$\frac{\sin(\pi n/2)}{\pi n} = \frac{1}{2\pi}\int_{-\pi/2}^{\pi/2}(1)e^{jn\omega}d\omega.$$

The right-hand side of this equation has the form of the Fourier transform synthesis equation for $x[n]$, where

$$X(e^{j\omega}) = \begin{cases} 1 & |\omega| \le \pi/2 \\ 0 & \pi/2 < |\omega| \le \pi \end{cases}.$$

In Table 5.3, we present a compact summary of the Fourier series and Fourier transform expressions for both continuous-time and discrete-time signals, and we also indicate the duality relationships that apply in each case.

TABLE 5.3 SUMMARY OF FOURIER SERIES AND TRANSFORM EXPRESSIONS

	Continuous time		Discrete time	
	Time domain	**Frequency domain**	**Time domain**	**Frequency domain**
Fourier Series	$x(t) =$ $\sum_{k=-\infty}^{+\infty} a_k e^{jk\omega_0 t}$ continuous time periodic in time	$a_k =$ $\frac{1}{T_0}\int_{T_0} x(t)e^{-jk\omega_0 t}$ discrete frequency aperiodic in frequency ⟺ duality	$x[n] =$ $\sum_{k=\langle N\rangle} a_k e^{jk(2\pi/N)n}$ discrete time periodic in time ⟺ duality	$a_k =$ $\frac{1}{N}\sum_{k=\langle N\rangle} x[n]e^{-jk(2\pi/N)n}$ discrete frequency periodic in frequency
Fourier Transform	$x(t) =$ $\frac{1}{2\pi}\int_{-\infty}^{+\infty} X(j\omega)e^{j\omega t}d\omega$ continuous time aperiodic in time ⟺ duality	$X(j\omega) =$ $\int_{-\infty}^{+\infty} x(t)e^{-j\omega t}dt$ **continuous frequency** **aperiodic in frequency**	$x[n] =$ $\frac{1}{2\pi}\int_{2\pi} X(e^{j\omega})e^{j\omega n}d\omega$ **discrete time** **aperiodic in time**	$X(e^{j\omega}) =$ $\sum_{n=-\infty}^{+\infty} x[n]e^{-j\omega n}$ continuous frequency periodic in frequency

5.8 SYSTEMS CHARACTERIZED BY LINEAR CONSTANT-COEFFICIENT DIFFERENCE EQUATIONS

A general linear constant-coefficient difference equation for an LTI system with input $x[n]$ and output $y[n]$ is of the form

$$\sum_{k=0}^{N} a_k y[n-k] = \sum_{k=0}^{M} b_k x[n-k]. \tag{5.78}$$

The class of systems described by such difference equations is quite an important and useful one. In this section, we take advantage of several of the properties of the discrete-time Fourier transform to determine the frequency response $H(e^{j\omega})$ for an LTI system described by such an equation. The approach we follow closely parallels the discussion in Section 4.7 for continuous-time LTI systems described by linear constant-coefficient differential equations.

There are two related ways in which to determine $H(e^{j\omega})$. The first of these, which we illustrated in Section 3.11 for several simple difference equations, explicitly uses the fact that complex exponentials are eigenfunctions of LTI systems. Specifically, if $x[n] = e^{j\omega n}$ is the input to an LTI system, then the output must be of the form $H(e^{j\omega})e^{j\omega n}$. Substituting these expressions into eq. (5.78) and performing some algebra allows us to solve for $H(e^{j\omega})$. In this section, we follow a second approach making use of the convolution, linearity, and time-shifting properties of the discrete-time Fourier transform. Let $X(e^{j\omega})$, $Y(e^{j\omega})$, and $H(e^{j\omega})$ denote the Fourier transforms of the input $x[n]$, output $y[n]$, and impulse response $h[n]$, respectively. The convolution property, eq. (5.48), of the discrete-time Fourier transform then implies that

$$H(e^{j\omega}) = \frac{Y(e^{j\omega})}{X(e^{j\omega})}. \tag{5.79}$$

Applying the Fourier transform to both sides of eq. (5.78) and using the linearity and time-shifting properties, we obtain the expression

$$\sum_{k=0}^{N} a_k e^{-jk\omega} Y(e^{j\omega}) = \sum_{k=0}^{M} b_k e^{-jk\omega} X(e^{j\omega}),$$

or equivalently,

$$H(e^{j\omega}) = \frac{Y(e^{j\omega})}{X(e^{j\omega})} = \frac{\sum_{k=0}^{M} b_k e^{-jk\omega}}{\sum_{k=0}^{N} a_k e^{-jk\omega}}. \tag{5.80}$$

Comparing eq. (5.80) with eq. (4.76), we see that, as in the case of continuous time, $H(e^{j\omega})$ is a ratio of polynomials, but in discrete time the polynomials are in the variable $e^{-j\omega}$. The coefficients of the numerator polynomial are the same coefficients as appear on the right side of eq. (5.78), and the coefficients of the denominator polynomial are the same as appear on the left side of that equation. Therefore, the frequency response of the LTI system specified by eq. (5.78) can be written down by inspection.

The difference equation (5.78) is generally referred to as an Nth-order difference equation, as it involves delays in the output $y[n]$ of up to N time steps. Also, the denominator of $H(e^{j\omega})$ in eq. (5.80) is an Nth-order polynomial in $e^{-j\omega}$.

Example 5.18

Consider the causal LTI system that is characterized by the difference equation

$$y[n] - ay[n-1] = x[n], \tag{5.81}$$

with $|a| < 1$. From eq. (5.80), the frequency response of this system is

$$H(e^{j\omega}) = \frac{1}{1 - ae^{-j\omega}}. \tag{5.82}$$

Comparing this with Example 5.1, we recognize it as the Fourier transform of the sequence $a^n u[n]$. Thus, the impulse response of the system is

$$h[n] = a^n u[n]. \tag{5.83}$$

Example 5.19

Consider a causal LTI system that is characterized by the difference equation

$$y[n] - \frac{3}{4}y[n-1] + \frac{1}{8}y[n-2] = 2x[n]. \tag{5.84}$$

From eq. (5.80), the frequency response is

$$H(e^{j\omega}) = \frac{2}{1 - \frac{3}{4}e^{-j\omega} + \frac{1}{8}e^{-j2\omega}}. \tag{5.85}$$

As a first step in obtaining the impulse response, we factor the denominator of eq. (5.85):

$$H(e^{j\omega}) = \frac{2}{(1 - \frac{1}{2}e^{-j\omega})(1 - \frac{1}{4}e^{-j\omega})}. \tag{5.86}$$

$H(e^{j\omega})$ can be expanded by the method of partial fractions, as in Example A.3 in the appendix. The result of this expansion is

$$H(e^{j\omega}) = \frac{4}{1 - \frac{1}{2}e^{-j\omega}} - \frac{2}{1 - \frac{1}{4}e^{-j\omega}}. \tag{5.87}$$

The inverse transform of each term can be recognized by inspection, with the result that

$$h[n] = 4\left(\frac{1}{2}\right)^n u[n] - 2\left(\frac{1}{4}\right)^n u[n]. \tag{5.88}$$

The procedure followed in Example 5.19 is identical in style to that used in continuous time. Specifically, after expanding $H(e^{j\omega})$ by the method of partial fractions, we can find the inverse transform of each term by inspection. The same approach can be applied to the frequency response of any LTI system described by a linear constant-coefficient difference equation in order to determine the system impulse response. Also, as illustrated in the next example, if the Fourier transform $X(e^{j\omega})$ of the input to such a system is a ratio of polynomials in $e^{-j\omega}$, then $Y(e^{j\omega})$ is as well. In this case, we can use the same technique to find the response $y[n]$ to the input $x[n]$.

Example 5.20

Consider the LTI system of Example 5.19, and let the input to this system be

$$x[n] = \left(\frac{1}{4}\right)^n u[n].$$

Then, using eq. (5.80) and Example 5.1 or 5.18, we obtain

$$\begin{aligned} Y(e^{j\omega}) = H(e^{j\omega})X(e^{j\omega}) &= \left[\frac{2}{(1-\frac{1}{2}e^{-j\omega})(1-\frac{1}{4}e^{-j\omega})}\right]\left[\frac{1}{1-\frac{1}{4}e^{-j\omega}}\right] \\ &= \frac{2}{(1-\frac{1}{2}e^{-j\omega})(1-\frac{1}{4}e^{-e\omega})^2}. \end{aligned} \tag{5.89}$$

As described in the appendix, the form of the partial-fraction expansion in this case is

$$Y(e^{j\omega}) = \frac{B_{11}}{1-\frac{1}{4}e^{-j\omega}} + \frac{B_{12}}{(1-\frac{1}{4}e^{-j\omega})^2} + \frac{B_{21}}{1-\frac{1}{2}e^{-j\omega}}, \tag{5.90}$$

where the constants B_{11}, B_{12}, and B_{21} can be determined using the techniques described in the appendix. This particular expansion is worked out in detail in Example A.4, and the values obtained are

$$B_{11} = -4, \quad B_{12} = -2, \quad B_{21} = 8,$$

so that

$$Y(e^{j\omega}) = -\frac{4}{1-\frac{1}{4}e^{-j\omega}} - \frac{2}{(1-\frac{1}{4}e^{-j\omega})^2} + \frac{8}{1-\frac{1}{2}e^{-j\omega}}. \tag{5.91}$$

The first and third terms are of the same type as those encountered in Example 5.19, while the second term is of the same form as one seen in Example 5.13. Either from these examples or from Table 5.2, we can invert each of the terms in eq. (5.91) to obtain the inverse transform

$$y[n] = \left\{-4\left(\frac{1}{4}\right)^n - 2(n+1)\left(\frac{1}{4}\right)^n + 8\left(\frac{1}{2}\right)^n\right\}u[n]. \tag{5.92}$$

5.9 SUMMARY

In this chapter, we have paralleled Chapter 4 as we developed the Fourier transform for discrete-time signals and examined many of its important properties. Throughout the chapter, we have seen a great many similarities between continuous-time and discrete-time Fourier analysis, and we have also seen some important differences. For example, the relationship between Fourier series and Fourier transforms in discrete time is exactly analogous to that in continuous time. In particular, our derivation of the discrete-time Fourier transform for aperiodic signals from the discrete-time Fourier series representations is very much the same as the corresponding continuous-time derivation. Furthermore, many of the properties of continuous-time transforms have exact discrete-time counterparts. On the other hand, in contrast to the continuous-time case, the discrete-time Fourier transform of an aperiodic signal is always periodic with period 2π. In addition to similarities and differences such as these, we have described the duality relationships among the Fourier representations of continuous-time and discrete-time signals.

The most important similiarities between continuous- and discrete-time Fourier analysis are in their uses in analyzing and representing signals and LTI systems. Specifically, the convolution property provides us with the basis for the frequency-domain analysis of LTI systems. We have already seen some of the utility of this approach in our discussion of filtering in Chapters 3–5 and in our examination of systems described by linear constant-coefficient differential or difference equations, and we will gain a further appreciation for its utility in Chapter 6, in which we examine filtering and time-versus-frequency issues in more detail. In addition, the multiplication properties in continuous and discrete time are essential to our development of sampling in Chapter 7 and communications in Chapter 8.

Chapter 5 Problems

The first section of problems belongs to the basic category and the answers are provided in the back of the book. The remaining three sections contain problems belonging to the basic, advanced, and extension categories, respectively.

BASIC PROBLEMS WITH ANSWERS

5.1. Use the Fourier transform analysis equation (5.9) to calculate the Fourier transforms of:

(a) $(\frac{1}{2})^{n-1}u[n-1]$ **(b)** $(\frac{1}{2})^{|n-1|}$

Sketch and label one period of the magnitude of each Fourier transform.

5.2. Use the Fourier transform analysis equation (5.9) to calculate the Fourier transforms of:

(a) $\delta[n-1]+\delta[n+1]$ **(b)** $\delta[n+2]-\delta[n-2]$

Sketch and label one period of the magnitude of each Fourier transform.

5.3. Determine the Fourier transform for $-\pi \leq \omega < \pi$ in the case of each of the following periodic signals:

(a) $\sin(\frac{\pi}{3}n+\frac{\pi}{4})$ **(b)** $2+\cos(\frac{\pi}{6}n+\frac{\pi}{8})$

5.4. Use the Fourier transform synthesis equation (5.8) to determine the inverse Fourier transforms of:

(a) $X_1(e^{j\omega}) = \sum_{k=-\infty}^{\infty}\{2\pi\delta(\omega-2\pi k)+\pi\delta(\omega-\frac{\pi}{2}-2\pi k)+\pi\delta(\omega+\frac{\pi}{2}-2\pi k)\}$

(b) $X_2(e^{j\omega}) = \begin{cases} 2j, & 0<\omega\leq\pi \\ -2j, & -\pi<\omega\leq 0 \end{cases}$

5.5. Use the Fourier transform synthesis equation (5.8) to determine the inverse Fourier transform of $X(e^{j\omega}) = |X(e^{j\omega})|e^{j\sphericalangle X(e^{j\omega})}$, where

$$|X(e^{j\omega})| = \begin{cases} 1, & 0\leq|\omega|<\frac{\pi}{4} \\ 0, & \frac{\pi}{4}\leq|\omega|\leq\pi \end{cases} \quad \text{and } \sphericalangle X(e^{j\omega}) = -\frac{3\omega}{2}.$$

Use your answer to determine the values of n for which $x[n] = 0$.

5.6. Given that $x[n]$ has Fourier transform $X(e^{j\omega})$, express the Fourier transforms of the following signals in terms of $X(e^{j\omega})$. You may use the Fourier transform properties listed in Table 5.1.

(a) $x_1[n] = x[1-n]+x[-1-n]$

(b) $x_2[n] = \frac{x^*[-n]+x[n]}{2}$

(c) $x_3[n] = (n-1)^2x[n]$

5.7. For each of the following Fourier transforms, use Fourier transform properties (Table 5.1) to determine whether the corresponding time-domain signal is (i) real, imaginary, or neither and (ii) even, odd, or neither. Do this without evaluating the inverse of any of the given transforms.

(a) $X_1(e^{j\omega}) = e^{-j\omega}\sum_{k=1}^{10}(\sin k\omega)$

(b) $X_2(e^{j\omega}) = j\sin(\omega)\cos(5\omega)$

(c) $X_3(e^{j\omega}) = A(\omega) + e^{jB(\omega)}$ where

$$A(\omega) = \begin{cases} 1, & 0 \leq |\omega| \leq \frac{\pi}{8} \\ 0, & \frac{\pi}{8} < |\omega| \leq \pi \end{cases} \quad \text{and } B(\omega) = -\frac{3\omega}{2} + \pi.$$

5.8. Use Tables 5.1 and 5.2 to help determine $x[n]$ when its Fourier transform is

$$X(e^{j\omega}) = \frac{1}{1-e^{-j\omega}}\left(\frac{\sin\frac{3}{2}\omega}{\sin\frac{\omega}{2}}\right) + 5\pi\delta(\omega), \qquad -\pi < \omega \leq \pi$$

5.9. The following four facts are given about a real signal $x[n]$ with Fourier transform $X(e^{jw})$:

1. $x[n] = 0$ for $n > 0$.
2. $x[0] > 0$.
3. $\mathcal{Im}\{X(e^{j\omega})\} = \sin\omega - \sin 2\omega$.
4. $\frac{1}{2\pi}\int_{-\pi}^{\pi}|X(e^{j\omega})|^2\,d\omega = 3$.

Determine $x[n]$.

5.10. Use Tables 5.1 and 5.2 in conjunction with the fact that

$$X(e^{j0}) = \sum_{n=-\infty}^{\infty} x[n]$$

to determine the numerical value of

$$A = \sum_{n=0}^{\infty} n\left(\frac{1}{2}\right)^n.$$

5.11. Consider a signal $g[n]$ with Fourier transform $G(e^{j\omega})$. Suppose

$$g[n] = x_{(2)}[n],$$

where the signal $x[n]$ has a Fourier transform $X(e^{j\omega})$. Determine a real number α such that $0 < \alpha < 2\pi$ and $G(e^{j\omega}) = G(e^{j(\omega-\alpha)})$.

5.12. Let

$$y[n] = \left(\frac{\sin\frac{\pi}{4}n}{\pi n}\right)^2 * \left(\frac{\sin\omega_c n}{\pi n}\right),$$

where $*$ denotes convolution and $|\omega_c| \leq \pi$. Determine a stricter constraint on ω_c

which ensures that

$$y[n] = \left(\frac{\sin \frac{\pi}{4} n}{\pi n}\right)^2.$$

5.13. An LTI system with impulse response $h_1[n] = (\frac{1}{3})^n u[n]$ is connected in parallel with another causal LTI system with impulse response $h_2[n]$. The resulting parallel interconnection has the frequency response

$$H(e^{j\omega}) = \frac{-12 + 5e^{-j\omega}}{12 - 7e^{-j\omega} + e^{-j2\omega}}.$$

Determine $h_2[n]$.

5.14. Suppose we are given the following facts about an LTI system S with impulse response $h[n]$ and frequency response $H(e^{j\omega})$:

1. $(\frac{1}{4})^n u[n] \longrightarrow g[n]$, where $g[n] = 0$ for $n \geq 2$ and $n < 0$.
2. $H(e^{j\pi/2}) = 1$.
3. $H(e^{j\omega}) = H(e^{j(\omega - \pi)})$.

Determine $h[n]$.

5.15. Let the inverse Fourier transform of $Y(e^{j\omega})$ be

$$y[n] = \left(\frac{\sin \omega_c n}{\pi n}\right)^2,$$

where $0 < \omega_c < \pi$. Determine the value of ω_c which ensures that

$$Y(e^{j\pi}) = \frac{1}{2}.$$

5.16. The Fourier transform of a particular signal is

$$X(e^{j\omega}) = \sum_{k=0}^{3} \frac{(1/2)^k}{1 - \frac{1}{4}e^{-j(\omega - \pi/2k)}}.$$

It can be shown that

$$x[n] = g[n]q[n],$$

where $g[n]$ is of the form $\alpha^n u[n]$ and $q[n]$ is a periodic signal with period N.

(a) Determine the value of α.
(b) Determine the value of N.
(c) Is $x[n]$ real?

5.17. The signal $x[n] = (-1)^n$ has a fundamental period of 2 and corresponding Fourier series coefficients a_k. Use duality to determine the Fourier series coefficients b_k of the signal $g[n] = a_n$ with a fundamental period of 2.

5.18. Given the fact that

$$a^{|n|} \overset{\mathfrak{F}}{\longleftrightarrow} \frac{1 - a^2}{1 - 2a\cos\omega + a^2}, \quad |a| < 1,$$

use duality to determine the Fourier series coefficients of the following continuous-time signal with period $T = 1$:

$$x(t) = \frac{1}{5 - 4\cos(2\pi t)}.$$

5.19. Consider a causal and stable LTI system S whose input $x[n]$ and output $y[n]$ are related through the second-order difference equation

$$y[n] - \frac{1}{6}y[n-1] - \frac{1}{6}y[n-2] = x[n].$$

(a) Determine the frequency response $H(e^{j\omega})$ for the system S.
(b) Determine the impulse response $h[n]$ for the system S.

5.20. A causal and stable LTI system S has the property that

$$\left(\frac{4}{5}\right)^n u[n] \longrightarrow n\left(\frac{4}{5}\right)^n u[n].$$

(a) Determine the frequency response $H(e^{j\omega})$ for the system S.
(b) Determine a difference equation relating any input $x[n]$ and the corresponding output $y[n]$.

BASIC PROBLEMS

5.21. Compute the Fourier transform of each of the following signals:
(a) $x[n] = u[n-2] - u[n-6]$
(b) $x[n] = (\frac{1}{2})^{-n}u[-n-1]$
(c) $x[n] = (\frac{1}{3})^{|n|}u[-n-2]$
(d) $x[n] = 2^n \sin(\frac{\pi}{4}n)u[-n]$
(e) $x[n] = (\frac{1}{2})^{|n|}\cos(\frac{\pi}{8}(n-1))$
(f) $x[n] = \begin{cases} n, & -3 \le n \le 3 \\ 0, & \text{otherwise} \end{cases}$
(g) $x[n] = \sin(\frac{\pi}{2}n) + \cos(n)$
(h) $x[n] = \sin(\frac{5\pi}{3}n) + \cos(\frac{7\pi}{3}n)$
(i) $x[n] = x[n-6]$, and $x[n] = u[n] - u[n-5]$ for $0 \le n \le 5$
(j) $x[n] = (n-1)(\frac{1}{3})^{|n|}$
(k) $x[n] = (\frac{\sin(\pi n/5)}{\pi n})\cos(\frac{7\pi}{2}n)$

5.22. The following are the Fourier transforms of discrete-time signals. Determine the signal corresponding to each transform.
(a) $X(e^{j\omega}) = \begin{cases} 1, & \frac{\pi}{4} \le |\omega| \le \frac{3\pi}{4} \\ 0, & \frac{3\pi}{4} \le |\omega| \le \pi, 0 \le |\omega| < \frac{\pi}{4} \end{cases}$
(b) $X(e^{j\omega}) = 1 + 3e^{-j\omega} + 2e^{-j2\omega} - 4e^{-j3\omega} + e^{-j10\omega}$
(c) $X(e^{j\omega}) = e^{-j\omega/2}$ for $-\pi \le \omega \le \pi$
(d) $X(e^{j\omega}) = \cos^2\omega + \sin^2 3\omega$

(e) $X(e^{j\omega}) = \sum_{k=-\infty}^{\infty}(-1)^k\delta(\omega - \frac{\pi}{2}k)$

(f) $X(e^{j\omega}) = \frac{e^{-j\omega} - \frac{1}{5}}{1 - \frac{1}{5}e^{-j\omega}}$

(g) $X(e^{j\omega}) = \frac{1 - \frac{1}{3}e^{-j\omega}}{1 - \frac{1}{4}e^{-j\omega} - \frac{1}{8}e^{-2j\omega}}$

(h) $X(e^{j\omega}) = \frac{1 - (\frac{1}{3})^6 e^{-j6\omega}}{1 - \frac{1}{3}e^{-j\omega}}$

5.23. Let $X(e^{j\omega})$ denote the Fourier transform of the signal $x[n]$ depicted in Figure P5.23. Perform the following calculations without explicitly evaluating $X(e^{j\omega})$:

(a) Evaluate $X(e^{j0})$.

(b) Find $\sphericalangle X(e^{j\omega})$.

(c) Evaluate $\int_{-\pi}^{\pi} X(e^{j\omega})d\omega$.

(d) Find $X(e^{j\pi})$.

(e) Determine and sketch the signal whose Fourier transform is $\mathcal{R}e\{x(\omega)\}$.

(f) Evaluate:

(i) $\int_{-\pi}^{\pi} |X(e^{j\omega})|^2 d\omega$

(ii) $\int_{-\pi}^{\pi} \left|\frac{dX(e^{j\omega})}{d\omega}\right|^2 d\omega$

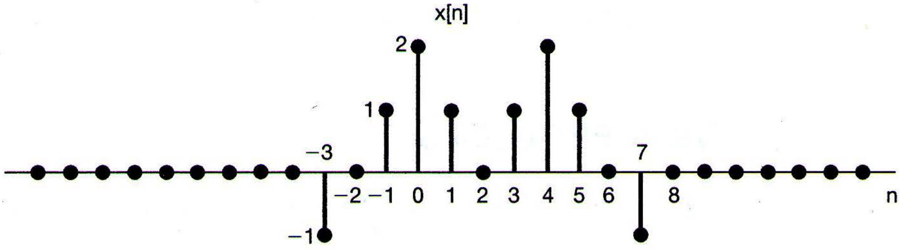


Fig P5.23

5.24. Determine which, if any, of the following signals have Fourier transforms that satisfy each of the following conditions:

1. $\mathcal{R}e\{X(e^{j\omega})\} = 0$.
2. $\mathcal{I}m\{X(e^{j\omega})\} = 0$.
3. There exists an integer α such that $e^{j\alpha\omega}X(e^{j\omega})$ is real.
4. $\int_{-\pi}^{\pi} X(e^{j\omega})d\omega = 0$.
5. $X(e^{j\omega})$ periodic.
6. $X(e^{j0}) = 0$.

(a) $x[n]$ as in Figure P5.24(a)

(b) $x[n]$ as in Figure P5.24(b)

(c) $x[n] = (\frac{1}{2})^n u[n]$

(d) $x[n] = (\frac{1}{2})^{|n|}$

(e) $x[n] = \delta[n-1] + \delta[n+2]$

(f) $x[n] = \delta[n-1] + \delta[n+3]$

(g) $x[n]$ as in Figure P5.24(c)

(h) $x[n]$ as in Figure P5.24(d)

(i) $x[n] = \delta[n-1] - \delta[n+1]$

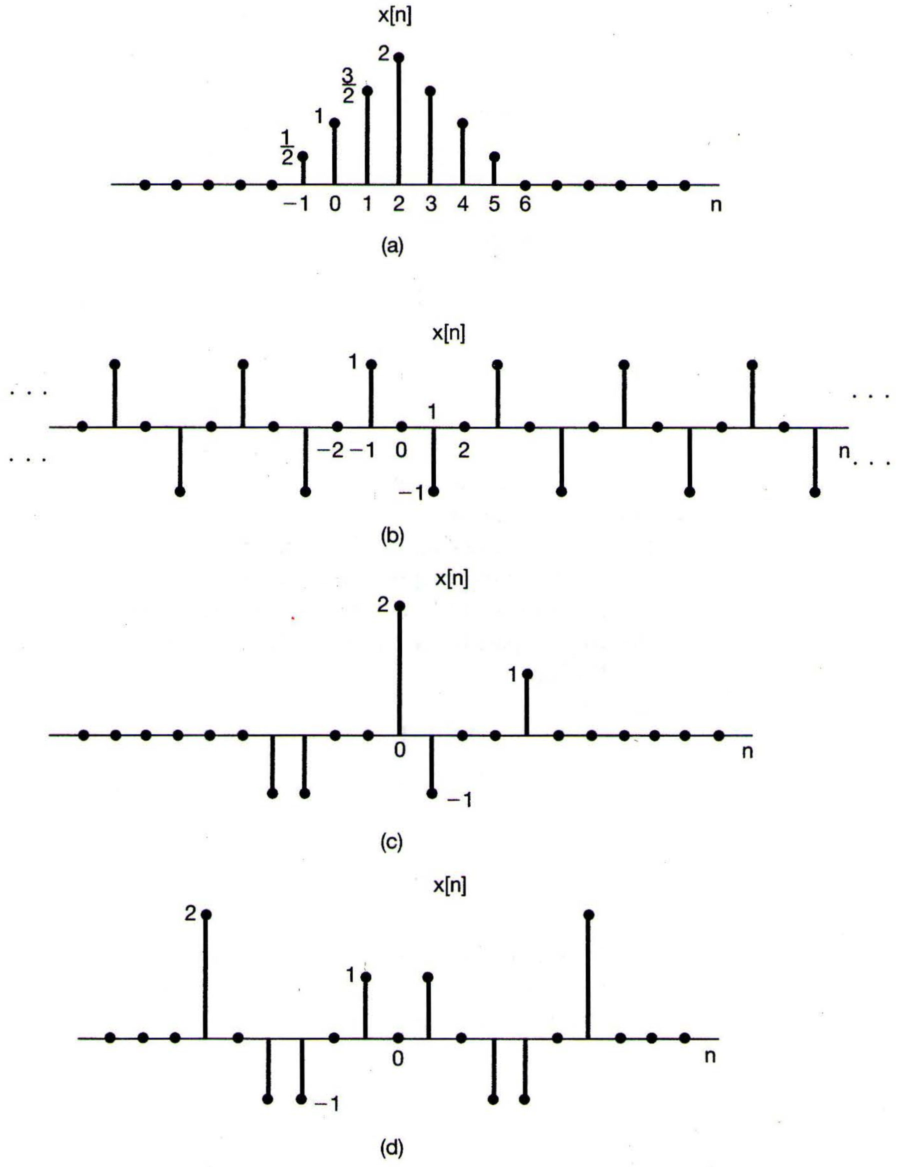


Fig P5.24

5.25. Consider the signal depicted in Figure P5.25. Let the Fourier transform of this signal be written in rectangular form as

$$X(e^{j\omega}) = A(\omega) + jB(\omega).$$

Sketch the function of time corresponding to the transform

$$Y(e^{j\omega}) = [B(\omega) + A(\omega)e^{j\omega}].$$

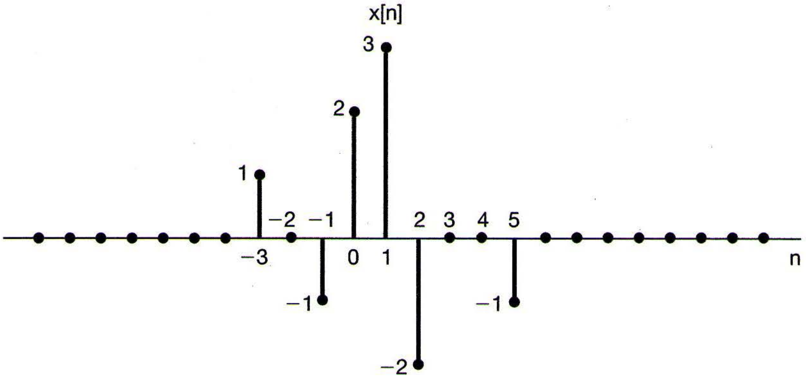


Fig P5.25

5.26. Let $x_1[n]$ be the discrete-time signal whose Fourier transform $X_1(e^{j\omega})$ is depicted in Figure P5.26(a).

(a) Consider the signal $x_2[n]$ with Fourier transform $X_2(e^{j\omega})$, as illustrated in Figure P5.26(b). Express $x_2[n]$ in terms of $x_1[n]$. [*Hint:* First express $X_2(e^{j\omega})$ in terms of $X_1(e^{j\omega})$, and then use properties of the Fourier transform.]

(b) Repeat part (a) for $x_3[n]$ with Fourier transform $X_3(e^{j\omega})$, as shown in Figure P5.26(c).

(c) Let

$$\alpha = \frac{\sum_{n=-\infty}^{\infty} n x_1[n]}{\sum_{n=-\infty}^{\infty} x_1[n]}.$$

This quantity, which is the center of gravity of the signal $x_1[n]$, is usually referred to as the *delay time* of $x_1[n]$. Find α. (You can do this without first determining $x_1[n]$ explicitly.)

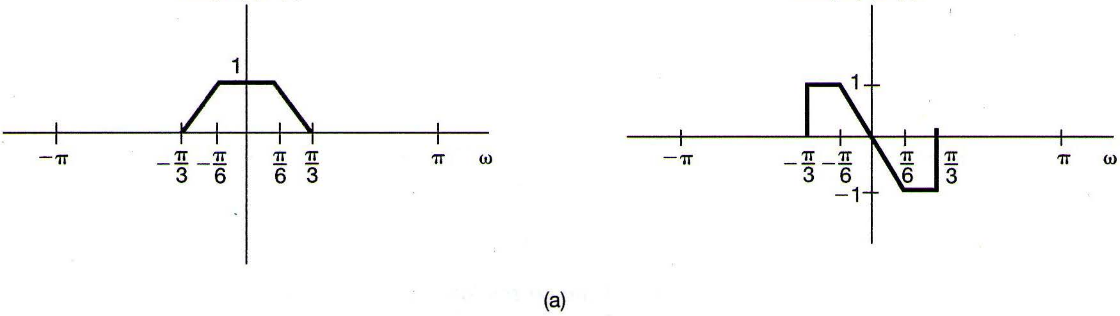
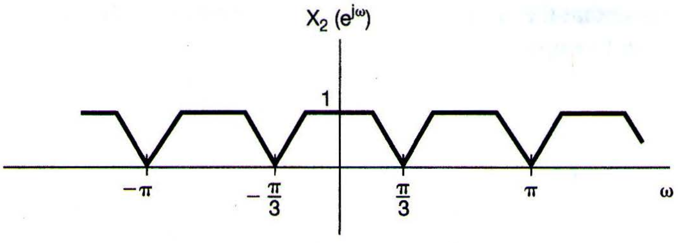
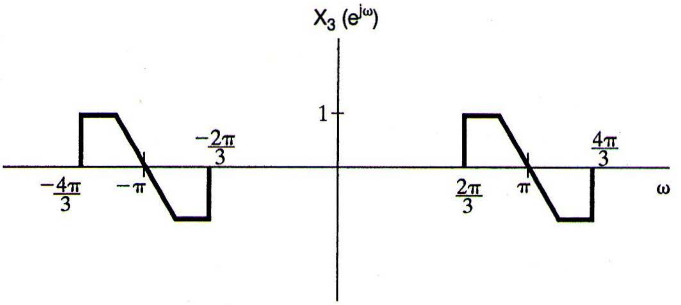


Fig P5.26a



(b)



(c)

Fig P5.26b,c

(d) Consider the signal $x_4[n] = x_1[n] * h[n]$, where

$$h[n] = \frac{\sin(\pi n/6)}{\pi n}.$$

Sketch $X_4(e^{j\omega})$.

5.27. (a) Let $x[n]$ be a discrete-time signal with Fourier transform $X(e^{j\omega})$, which is illustrated in Figure P5.27. Sketch the Fourier transform of

$$w[n] = x[n]p[n]$$

for each of the following signals $p[n]$:

(i) $p[n] = \cos \pi n$

(ii) $p[n] = \cos(\pi n/2)$

(iii) $p[n] = \sin(\pi n/2)$

(iv) $p[n] = \sum_{k=-\infty}^{\infty} \delta[n-2k]$

(v) $p[n] = \sum_{k=-\infty}^{\infty} \delta[n-4k]$

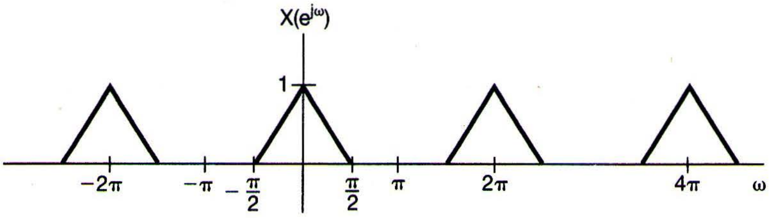


Fig P5.27

(b) Suppose that the signal $w[n]$ of part (a) is applied as the input to an LTI system with unit sample response

$$h[n] = \frac{\sin(\pi n/2)}{\pi n}.$$

Determine the output $y[n]$ for each of the choices of $p[n]$ in part (a).

5.28. The signals $x[n]$ and $g[n]$ are known to have Fourier transforms $X(e^{j\omega})$ and $G(e^{j\omega})$, respectively. Furthermore, $X(e^{j\omega})$ and $G(e^{j\omega})$ are related as follows:

$$\frac{1}{2\pi}\int_{-\pi}^{+\pi} X(e^{j\theta})G(e^{j(\omega-\theta)})d\theta = 1 + e^{-j\omega} \tag{P5.28–1}$$

(a) If $x[n] = (-1)^n$, determine a sequence $g[n]$ such that its Fourier transform $G(e^{j\omega})$ satisfies eq. (P5.28–1). Are there other possible solutions for $g[n]$?

(b) Repeat the previous part for $x[n] = (\frac{1}{2})^n u[n]$.

5.29. **(a)** Consider a discrete-time LTI system with impulse response

$$h[n] = \left(\frac{1}{2}\right)^n u[n].$$

Use Fourier transforms to determine the response to each of the following input signals:

(i) $x[n] = (\frac{3}{4})^n u[n]$
(ii) $x[n] = (n+1)(\frac{1}{4})^n u[n]$
(iii) $x[n] = (-1)^n$

(b) Suppose that

$$h[n] = \left[\left(\frac{1}{2}\right)^n \cos\left(\frac{\pi n}{2}\right)\right] u[n].$$

Use Fourier transforms to determine the response to each of the following inputs:

(i) $x[n] = (\frac{1}{2})^n u[n]$
(ii) $x[n] = \cos(\pi n/2)$

(c) Let $x[n]$ and $h[n]$ be signals with the following Fourier transforms:

$$X(e^{j\omega}) = 3e^{j\omega} + 1 - e^{-j\omega} + 2e^{-j3\omega},$$
$$H(e^{j\omega}) = -e^{j\omega} + 2e^{-2j\omega} + e^{j4\omega}.$$

Determine $y[n] = x[n] * h[n]$.

5.30. In Chapter 4, we indicated that the continuous-time LTI system with impulse response

$$h(t) = \frac{W}{\pi}\operatorname{sinc}\left(\frac{Wt}{\pi}\right) = \frac{\sin Wt}{\pi t}$$

plays a very important role in LTI system analysis. The same is true of the discrete-time LTI system with impulse response

$$h[n] = \frac{W}{\pi}\operatorname{sinc}\left(\frac{Wn}{\pi}\right) = \frac{\sin Wn}{\pi n}.$$

(a) Determine and sketch the frequency response for the system with impulse response $h[n]$.

(b) Consider the signal

$$x[n] = \sin\left(\frac{\pi n}{8}\right) - 2\cos\left(\frac{\pi n}{4}\right).$$

Suppose that this signal is the input to LTI systems with the following impulse responses. Determine the output in each case.

(i) $h[n] = \frac{\sin(\pi n/6)}{\pi n}$

(ii) $h[n] = \frac{\sin(\pi n/6)}{\pi n} + \frac{\sin(\pi n/2)}{\pi n}$

(iii) $h[n] = \frac{\sin(\pi n/6)\sin(\pi n/3)}{\pi^2 n^2}$

(iv) $h[n] = \frac{\sin(\pi n/6)\sin(\pi n/3)}{\pi n}$

(c) Consider an LTI system with unit sample response

$$h[n] = \frac{\sin(\pi n/3)}{\pi n}.$$

Determine the output for each of the following inputs:

(i) $x[n] =$ the square wave depicted in Figure P5.30

(ii) $x[n] = \sum_{k=-\infty}^{\infty} \delta[n - 8k]$

(iii) $x[n] = (-1)^n$ times the square wave depicted in Figure P5.30

(iv) $x[n] = \delta[n + 1] + \delta[n - 1]$

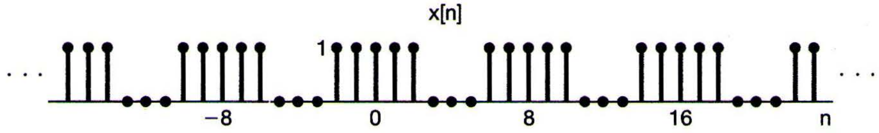


Fig P5.30

5.31. An LTI system S with impulse response $h[n]$ and frequency response $H(e^{j\omega})$ is known to have the property that, when $-\pi \le \omega_0 \le \pi$,

$$\cos\omega_0 n \longrightarrow \omega_0 \cos\omega_0 n.$$

(a) Determine $H(e^{j\omega})$.

(b) Determine $h[n]$.

5.32. Let $h_1[n]$ and $h_2[n]$ be the impulse responses of causal LTI systems, and let $H_1(e^{j\omega})$ and $H_2(e^{j\omega})$ be the corresponding frequency responses. Under these conditions, is the following equation true in general or not? Justify your answer.

$$\left[\frac{1}{2\pi}\int_{-\pi}^{\pi} H_1(e^{j\omega})d\omega\right]\left[\frac{1}{2\pi}\int_{-\pi}^{\pi} H_2(e^{j\omega})d\omega\right] = \frac{1}{2\pi}\int_{-\pi}^{\pi} H_1(e^{j\omega})H_2(e^{j\omega})d\omega.$$

5.33. Consider a causal LTI system described by the difference equation

$$y[n] + \frac{1}{2}y[n-1] = x[n].$$

(a) Determine the frequency response $H(e^{j\omega})$ of this system.
(b) What is the response of the system to the following inputs?
  (i) $x[n] = (\frac{1}{2})^n u[n]$
  (ii) $x[n] = (-\frac{1}{2})^n u[n]$
  (iii) $x[n] = \delta[n] + \frac{1}{2}\delta[n-1]$
  (iv) $x[n] = \delta[n] - \frac{1}{2}\delta[n-1]$
(c) Find the response to the inputs with the following Fourier transforms:
  (i) $X(e^{j\omega}) = \frac{1-\frac{1}{4}e^{-j\omega}}{1+\frac{1}{2}e^{-j\omega}}$
  (ii) $X(e^{j\omega}) = \frac{1+\frac{1}{2}e^{-j\omega}}{1-\frac{1}{4}e^{-j\omega}}$
  (iii) $X(e^{j\omega}) = \frac{1}{(1-\frac{1}{4}e^{-j\omega})(1+\frac{1}{2}e^{-j\omega})}$
  (iv) $X(e^{j\omega}) = 1 + 2e^{-3j\omega}$

5.34. Consider a system consisting of the cascade of two LTI systems with frequency responses

$$H_1(e^{j\omega}) = \frac{2 - e^{-j\omega}}{1 + \frac{1}{2}e^{-j\omega}}$$

and

$$H_2(e^{j\omega}) = \frac{1}{1 - \frac{1}{2}e^{-j\omega} + \frac{1}{4}e^{-j2\omega}}.$$

(a) Find the difference equation describing the overall system.
(b) Determine the impulse response of the overall system.

5.35. A causal LTI system is described by the difference equation

$$y[n] - ay[n-1] = bx[n] + x[n-1],$$

where a is real and less than 1 in magnitude.

(a) Find a value of b such that the frequency response of the system satisfies

$$|H(e^{j\omega})| = 1, \text{ for all } \omega.$$

This kind of system is called an *all-pass system*, as it does not attenuate the input $e^{j\omega n}$ for *any* value of ω. Use the value of b that you have found in the rest of the problem.

(b) Roughly sketch $\sphericalangle H(e^{j\omega})$, $0 \le \omega \le \pi$, when $a = \frac{1}{2}$.
(c) Roughly sketch $\sphericalangle H(e^{j\omega})$, $0 \le \omega \le \pi$, when $a = -\frac{1}{2}$.

(d) Find and plot the output of this system with $a = -\frac{1}{2}$ when the input is

$$x[n] = \left(\frac{1}{2}\right)^n u[n].$$

From this example, we see that a nonlinear change in phase can have a significantly different effect on a signal than the time shift that results from a linear phase.

5.36. **(a)** Let $h[n]$ and $g[n]$ be the impulse responses of two stable discrete-time LTI systems that are inverses of each other. What is the relationship between the frequency responses of these two systems?

(b) Consider causal LTI systems described by the following difference equations. In each case, determine the impulse response of the inverse system and the difference equation that characterizes the inverse.

(i) $y[n] = x[n] - \frac{1}{4}x[n-1]$

(ii) $y[n] + \frac{1}{2}y[n-1] = x[n]$

(iii) $y[n] + \frac{1}{2}y[n-1] = x[n] - \frac{1}{4}x[n-1]$

(iv) $y[n] + \frac{5}{4}y[n-1] - \frac{1}{8}y[n-2] = x[n] - \frac{1}{4}x[n-1] - \frac{1}{8}x[n-2]$

(v) $y[n] + \frac{5}{4}y[n-1] - \frac{1}{8}y[n-2] = x[n] - \frac{1}{2}x[n-1]$

(vi) $y[n] + \frac{5}{4}y[n-1] - \frac{1}{8}y[n-2] = x[n]$

(c) Consider the causal, discrete-time LTI system described by the difference equation

$$y[n] + y[n-1] + \frac{1}{4}y[n-2] = x[n-1] - \frac{1}{2}x[n-2]. \qquad \text{(P5.36–1)}$$

What is the inverse of this system? Show that the inverse is not causal. Find another causal LTI system that is an "inverse with delay" of the system described by eq. (P5.36–1). Specifically, find a causal LTI system such that the output $w[n]$ in Figure P5.36 equals $x[n-1]$.

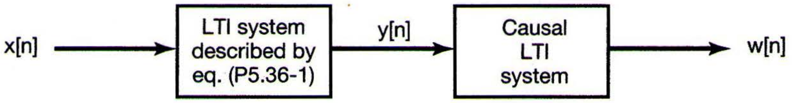


Fig P5.36

ADVANCED PROBLEMS

5.37. Let $X(e^{j\omega})$ be the Fourier transform of $x[n]$. Derive expressions in terms of $X(e^{j\omega})$ for the Fourier transforms of the following signals. (Do not assume that $x[n]$ is real.)

(a) $\mathcal{R}e\{x[n]\}$

(b) $x^*[-n]$

(c) $\mathcal{E}v\{x[n]\}$

5.38. Let $X(e^{j\omega})$ be the Fourier transform of a real signal $x[n]$. Show that $x[n]$ can be written as

$$x[n] = \int_0^{\pi} \{B(\omega)\cos\omega + C(\omega)\sin\omega\}\,d\omega$$

by finding expressions for $B(\omega)$ and $C(\omega)$ in terms of $X(e^{j\omega})$.

5.39. Derive the convolution property

$$x[n] * h[n] \overset{\mathcal{F}}{\longleftrightarrow} X(e^{j\omega})H(e^{j\omega}).$$

5.40. Let $x[n]$ and $h[n]$ be two signals, and let $y[n] = x[n] * h[n]$. Write two expressions for $y[0]$, one (using the convolution sum directly) in terms of $x[n]$ and $h[n]$, and one (using the convolution property of Fourier transforms) in terms of $X(e^{j\omega})$ and $H(e^{j\omega})$. Then, by a judicious choice of $h[n]$, use these two expressions to derive Parseval's relation—that is,

$$\sum_{n=-\infty}^{+\infty} |x[n]|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |X(e^{j\omega})|^2 d\omega.$$

In a similar fashion, derive the following generalization of Parseval's relation:

$$\sum_{n-\infty}^{+\infty} x[n]z^*[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} X(e^{j\omega})Z^*(e^{j\omega})d\omega.$$

5.41 Let $\tilde{x}[n]$ be a periodic signal with period N. A finite-duration signal $x[n]$ is related to $\tilde{x}[n]$ through

$$x[n] = \begin{cases} \tilde{x}[n], & n_0 \le n \le n_0 + N - 1 \\ 0, & \text{otherwise} \end{cases},$$

for some integer n_0. That is, $x[n]$ is equal to $\tilde{x}[n]$ over one period and zero elsewhere.

(a) If $\tilde{x}[n]$ has Fourier series coefficients a_k and $x[n]$ has Fourier transform $X(e^{j\omega})$, show that

$$a_k = \frac{1}{N}X(e^{j2\pi k/N})$$

regardless of the value of n_0.

(b) Consider the following two signals:

$$x[n] = u[n] - u[n-5]$$

$$\tilde{x}[n] = \sum_{k=-\infty}^{\infty} x[n-kN]$$

where N is a positive integer. Let a_k denote the Fourier coefficients of $\tilde{x}[n]$ and let $X(e^{j\omega})$ denote the Fourier transform of $x[n]$.

(i) Determine a closed-form expression for $X(e^{j\omega})$.

(ii) Using the result of part (i), determine an expression for the Fourier coefficients a_k.

5.42. In this problem, we derive the frequency-shift property of the discrete-time Fourier transform as a special case of the multiplication property. Let $x[n]$ be any discrete-time signal with Fourier transform $X(e^{j\omega})$, and let

$$g[n] = e^{j\omega_0 n}x[n].$$

(a) Determine and sketch the Fourier transform of

$$p[n] = e^{j\omega_0 n}.$$

(b) The multiplication property of the Fourier transform tells us that, since

$$g[n] = p[n]x[n],$$

$$G(e^{j\omega}) = \frac{1}{2\pi}\int_{<2\pi>} X(e^{j\theta})P(e^{j(\omega-\theta)})d\theta.$$

Evaluate this integral to show that

$$G(e^{j\omega}) = X(e^{j(\omega-\omega_0)}).$$

5.43. Let $x[n]$ be a signal with Fourier transform $X(e^{j\omega})$, and let

$$g[n] = x[2n]$$

be a signal whose Fourier transform is $G(e^{j\omega})$. In this problem, we derive the relationship between $G(e^{j\omega})$ and $X(e^{j\omega})$.

(a) Let

$$v[n] = \frac{(e^{-j\pi n}x[n]) + x[n]}{2}.$$

Express the Fourier transform $V(e^{j\omega})$ of $v[n]$ in terms of $X(e^{j\omega})$.

(b) Noting that $v[n] = 0$ for n odd, show that the Fourier transform of $v[2n]$ is equal to $V(e^{j\frac{\omega}{2}})$.

(c) Show that

$$x[2n] = v[2n].$$

It follows that

$$G(e^{j\omega}) = V(e^{j\omega/2}).$$

Now use the result of part (a) to express $G(e^{j\omega})$ in terms of $X(e^{j\omega})$.

5.44. **(a)** Let

$$x_1[n] = \cos\left(\frac{\pi n}{3}\right) + \sin\left(\frac{\pi n}{2}\right)$$

be a signal, and let $X_1(e^{j\omega})$ denote the Fourier transform of $x_1[n]$. Sketch $x_1[n]$, together with the signals with the following Fourier transforms:

(i) $X_2(e^{j\omega}) = X_1(e^{j\omega})e^{j\omega}, \ |\omega| < \pi$

(ii) $X_3(e^{j\omega}) = X_1(e^{j\omega})e^{-j3\omega/2}, \ |\omega| < \pi$

(b) Let

$$w(t) = \cos\left(\frac{\pi t}{3T}\right) + \sin\left(\frac{\pi t}{2T}\right)$$

be a continuous-time signal. Note that $x_1[n]$ can be regarded as a sequence of evenly spaced samples of $w(t)$; that is,

$$x_1[n] = w(nT).$$

Show that

$$x_2[n] = w(nT - \alpha)$$

and

$$x_3[n] = w(nT - \beta)$$

and specify the values of α and β. From this result we can conclude that $x_2[n]$ and $x_3[n]$ are also evenly spaced samples of $w(t)$.

5.45. Consider a discrete-time signal $x[n]$ with Fourier transform as illustrated in Figure P5.45. Provide dimensioned sketches of the following continuous-time signals:

(a) $x_1(t) = \sum_{n=-\infty}^{\infty} x[n]e^{j(2\pi/10)nt}$

(b) $x_2(t) = \sum_{n=-\infty}^{\infty} x[-n]e^{j(2\pi/10)nt}$

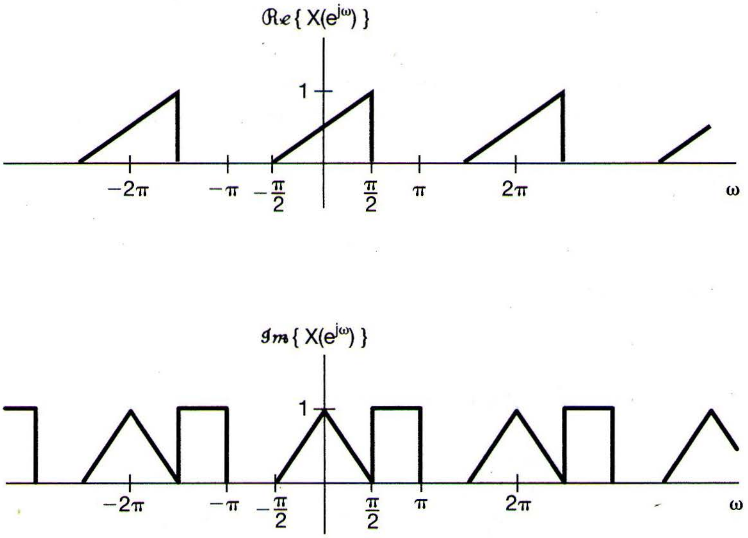


Fig P5.45

(c) $x_3(t) = \sum_{n=-\infty}^{\infty} \mathcal{O}d\{x[n]\}e^{j(2\pi/8)nt}$
(d) $x_4(t) = \sum_{n=-\infty}^{\infty} \mathcal{R}e\{x[n]\}e^{j(2\pi/6)nt}$

5.46. In Example 5.1, we showed that for $|\alpha| < 1$,

$$\alpha^n u[n] \overset{\mathcal{F}}{\longleftrightarrow} \frac{1}{1-\alpha e^{-j\omega}}.$$

(a) Use properties of the Fourier transform to show that

$$(n+1)\alpha^n u[n] \overset{\mathcal{F}}{\longleftrightarrow} \frac{1}{(1-\alpha e^{-j\omega})^2}.$$

(b) Show by induction that the inverse Fourier transform of

$$X(e^{j\omega}) = \frac{1}{(1-\alpha e^{-j\omega})^r}$$

is

$$x[n] = \frac{(n+r-1)!}{n!(r-1)!}\alpha^n u[n].$$

5.47. Determine whether each of the following statements is true or false. Justify your answers. In each statement, the Fourier transform of $x[n]$ is denoted by $X(e^{j\omega})$.
(a) If $X(e^{j\omega}) = X(e^{j(\omega-1)})$, then $x[n] = 0$ for $|n| > 0$.
(b) If $X(e^{j\omega}) = X(e^{j(\omega-\pi)})$, then $x[n] = 0$ for $|n| > 0$.
(c) If $X(e^{j\omega}) = X(e^{j\omega/2})$, then $x[n] = 0$ for $|n| > 0$.
(d) If $X(e^{j\omega}) = X(e^{j2\omega})$, then $x[n] = 0$ for $|n| > 0$.

5.48. We are given a discrete-time, linear, time-invariant, causal system with input denoted by $x[n]$ and output denoted by $y[n]$. This system is specified by the following *pair* of difference equations, involving an intermediate signal $w[n]$:

$$y[n] + \frac{1}{4}y[n-1] + w[n] + \frac{1}{2}w[n-1] = \frac{2}{3}x[n],$$

$$y[n] - \frac{5}{4}y[n-1] + 2w[n] - 2w[n-1] = -\frac{5}{3}x[n].$$

(a) Find the frequency response and unit sample response of the system.
(b) Find a single difference equation relating $x[n]$ and $y[n]$ for the system.

5.49. **(a)** A particular discrete-time system has input $x[n]$ and output $y[n]$. The Fourier transforms of these signals are related by the equation

$$Y(e^{j\omega}) = 2X(e^{j\omega}) + e^{-j\omega}X(e^{j\omega}) - \frac{dX(e^{j\omega})}{d\omega}.$$

(i) Is the system linear? Clearly justify your answer.
(ii) Is the system time invariant? Clearly justify your answer.
(iii) What is $y[n]$ if $x[n] = \delta[n]$?

(b) Consider a discrete-time system for which the transform $Y(e^{j\omega})$ of the output is related to the transform of the input through the relation

$$Y(e^{j\omega}) = \int_{\omega-\pi/4}^{\omega+\pi/4} X(e^{j\omega})d\omega.$$

Find an expression for $y[n]$ in terms of $x[n]$.

5.50. **(a)** Suppose we want to design a discrete-time LTI system which has the property that if the input is

$$x[n] = \left(\frac{1}{2}\right)^n u[n] - \frac{1}{4}\left(\frac{1}{2}\right)^{n-1} u[n-1],$$

then the output is

$$y[n] = \left(\frac{1}{3}\right)^n u[n].$$

(i) Find the impulse response *and* frequency response of a discrete-time LTI system that has the foregoing property.

(ii) Find a difference equation relating $x[n]$ and $y[n]$ that characterizes the system.

(b) Suppose that a system has the response $(1/4)^n u[n]$ to the input $(n+2)(1/2)^n u[n]$. If the output of this system is $\delta[n] - (-1/2)^n u[n]$, what is the input?

5.51. **(a)** Consider a discrete-time system with unit sample response

$$h[n] = \left(\frac{1}{2}\right)^n u[n] + \frac{1}{2}\left(\frac{1}{4}\right)^n u[n].$$

Determine a linear constant-coefficient difference equation relating the input and output of the system.

(b) Figure P5.51 depicts a block diagram implementation of a causal LTI system.

(i) Find a difference equation relating $x[n]$ and $y[n]$ for this system.

(ii) What is the frequency response of the system?

(iii) Determine the system's impulse response.

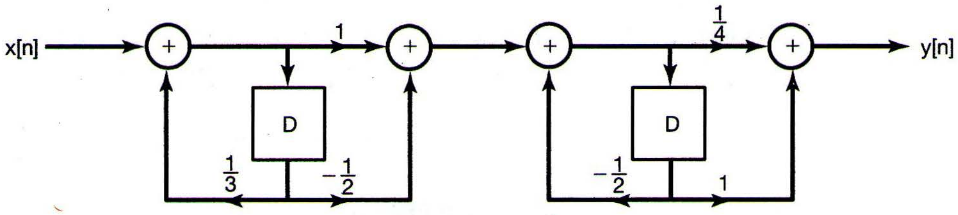


Fig P5.51

5.52. **(a)** Let $h[n]$ be the impulse response of a real, causal, discrete-time LTI system. Show that the system is completely specified by the real part of its frequency response. (*Hint:* Show how $h[n]$ can be recovered from $\mathcal{E}v\{h[n]\}$. What is the Fourier transform of $\mathcal{E}v\{h[n]\}$?) This is the discrete-time counterpart of the *real-part sufficiency* property of causal LTI systems considered in Problem 4.47 for continuous-time systems.

(b) Let $h[n]$ be real and causal. If

$$\mathcal{R}e\{H(e^{j\omega})\} = 1 + \alpha \cos 2\omega (\alpha \text{ real}),$$

determine $h[n]$ and $H(e^{j\omega})$.

(c) Show that $h[n]$ can be completely recovered from knowledge of $\mathcal{I}m\{H(e^{j\omega})\}$ and $h[0]$.

(d) Find two real, causal LTI systems whose frequency responses have imaginary parts equal to $\sin \omega$.

EXTENSION PROBLEMS

5.53. One of the reasons for the tremendous growth in the use of discrete-time methods for the analysis and synthesis of signals and systems was the development of exceedingly efficient tools for performing Fourier analysis of discrete-time sequences. At the heart of these methods is a technique that is very closely allied with discrete-time Fourier analysis and that is ideally suited for use on a digital computer or for implementation in digital hardware. This technique is the *discrete Fourier transform (DFT)* for finite-duration signals.

Let $x[n]$ be a signal of finite duration; that is, there is an integer N_1 so that

$$x[n] = 0, \qquad \text{outside the interval } 0 \le n \le N_1 - 1$$

Furthermore, let $X(e^{j\omega})$ denote the Fourier transform of $x[n]$. We can construct a periodic signal $\tilde{x}[n]$ that is equal to $x[n]$ over one period. Specifically, let $N \ge N_1$ be a given integer, and let $\tilde{x}[n]$ be periodic with period N and such that

$$\tilde{x}[n] = x[n], \qquad 0 \le n \le N - 1$$

The Fourier series coefficients for $\tilde{x}[n]$ are given by

$$a_k = \frac{1}{N} \sum_{\langle N \rangle} \tilde{x}[n] e^{-jk(2\pi/N)n}$$

Choosing the interval of summation to be that over which $\tilde{x}[n] = x[n]$, we obtain

$$a_k = \frac{1}{N} \sum_{n=0}^{N-1} x[n] e^{-jk(2\pi/N)n} \qquad \text{(P5.53–1)}$$

The set of coefficients defined by eq. (P5.53–1) comprise the DFT of $x[n]$. Specifically, the DFT of $x[n]$ is usually denoted by $\tilde{X}[k]$, and is defined as

$$\tilde{X}[k] = a_k = \frac{1}{N}\sum_{n=0}^{N-1} x[n]e^{-jk(2\pi/N)n}, \qquad k = 0, 1, \ldots, N-1 \qquad \text{(P5.53–2)}$$

The importance of the DFT stems from several facts. First note that the original finite duration signal can be recovered from its DFT. Specifically, we have

$$x[n] = \sum_{k=0}^{N-1} \tilde{X}[k]e^{jk(2\pi/N)n}, \qquad n = 0, 1, \ldots, N-1 \qquad \text{(P5.53–3)}$$

Thus, the finite-duration signal can either be thought of as being specified by the finite set of nonzero values it assumes or by the finite set of values of $\tilde{X}[k]$ in its DFT. A second important feature of the DFT is that there is an extremely fast algorithm, called the *fast Fourier transform (FFT),* for its calculation (see Problem 5.54 for an introduction to this extremely important technique). Also, because of its close relationship to the discrete-time Fourier series and transform, the DFT inherits some of their important properties.

(a) Assume that $N \geq N_1$. Show that

$$\tilde{X}[k] = \frac{1}{N}X\left(e^{j(2\pi k/N)}\right)$$

where $\tilde{X}[k]$ is the DFT of $x[n]$. That is, the DFT corresponds to samples of $X(e^{j\omega})$ taken every $2\pi/N$. Equation (P5.53–3) leads us to conclude that $x[n]$ can be uniquely represented by these samples of $X(e^{j\omega})$.

(b) Let us consider samples of $X(e^{j\omega})$ taken every $2\pi/M$, where $M < N_1$. These samples correspond to more than one sequence of duration N_1. To illustrate this, consider the two signals $x_1[n]$ and $x_2[n]$ depicted in Figure P5.53. Show that if we choose $M = 4$, we have

$$X_1\left(e^{j(2\pi k/4)}\right) = X_2\left(e^{j(2\pi k/4)}\right)$$

for all values of k.

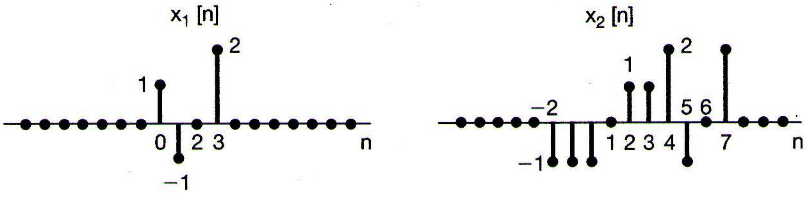


Fig P5.53

5.54. As indicated in Problem 5.53, there are many problems of practical importance in which one wishes to calculate the discrete Fourier transform (DFT) of discrete-time signals. Often, these signals are of quite long duration, and in such cases it is very important to use computationally efficient procedures. One of the reasons for the significant increase in the use of computerized techniques for the analysis of signals was the development of a very efficient technique known as the fast Fourier transform (FFT) algorithm for the calculation of the DFT of finite-duration sequences. In this problem, we develop the principle on which the FFT is based.

Let $x[n]$ be a signal that is 0 outside the interval $0 \leq n \leq N_1 - 1$. For $N \geq N_1$, the N-point DFT of $x[n]$ is given by

$$\tilde{X}[k] = \frac{1}{N} \sum_{k=0}^{N-1} x[n] e^{-jk(2\pi/N)n}, \quad k = 0, 1, \ldots, N-1. \qquad \text{(P5.54–1)}$$

It is convenient to write eq. (P5.54–1) as

$$\tilde{X}[k] = \frac{1}{N} \sum_{k=0}^{N-1} x[n] W_N^{nk}, \qquad \text{(P5.54–2)}$$

where

$$W_N = e^{-j2\pi/N}.$$

(a) One method for calculating $\tilde{X}[k]$ is by direct evaluation of eq. (P5.54–2). A useful measure of the complexity of such a computation is the total number of complex multiplications required. Show that the number of complex multiplications required to evaluate eq. (P5.54–2) directly, for $k = 0, 1, \ldots, N-1$, is N^2. Assume that $x[n]$ is complex and that the required values of W_N^{nk} have been precomputed and stored in a table. For simplicity, *do not* exploit the fact that, for certain values of n and k, W_N^{nk} is equal to ± 1 or $\pm j$ and hence does not, strictly speaking, require a full complex multiplication.

(b) Suppose that N is even. Let $f[n] = x[2n]$ represent the even-indexed samples of $x[n]$, and let $g[n] = x[2n+1]$ represent the odd-indexed samples.

(i) Show that $f[n]$ and $g[n]$ are zero outside the interval $0 \leq n \leq (N/2) - 1$.

(ii) Show that the N-point DFT $\tilde{X}[k]$ of $x[n]$ can be expressed as

$$\begin{aligned} \tilde{X}[k] &= \frac{1}{N} \sum_{n=0}^{(N/2)-1} f[n] W_{N/2}^{nk} + \frac{1}{N} W_N^k \sum_{n=0}^{(N/2)-1} g[n] W_{N/2}^{nk} \\ &= \frac{1}{2} \tilde{F}[k] + \frac{1}{2} W_N^k \tilde{G}[k], \quad k = 0, 1, \ldots, N-1, \qquad \text{(P5.54–3)} \end{aligned}$$

where

$$\tilde{F}[k] = \frac{2}{N} \sum_{n=0}^{(N/2)-1} f[n] W_{N/2}^{nk},$$

$$\tilde{G}[k] = \frac{2}{N} \sum_{n=0}^{(N/2)-1} g[n] W_{N/2}^{nk}.$$

(iii) Show that, for all k,

$$\tilde{F}\left[k + \frac{N}{2}\right] = \tilde{F}[k],$$

$$\tilde{G}\left[k + \frac{N}{2}\right] = \tilde{G}[k].$$

Note that $\tilde{F}[k]$, $k = 0, 1, \ldots, (N/2) - 1$, and $\tilde{G}[k]$, $k = 0, 1, \ldots, (N/2) - 1$, are the $(N/2)$-point DFTs of $f[n]$ and $g[n]$, respectively. Thus, eq. (P5.54–3) indicates that the length-N DFT of $x[n]$ can be calculated in terms of two DFTs of length $N/2$.

(iv) Determine the number of complex multiplications required to compute $\tilde{X}[k]$, $k = 0, 1, 2, \ldots, N - 1$, from eq. (P5.54–3) by first computing $\tilde{F}[k]$ and $\tilde{G}[k]$. [Make the same assumptions about multiplications as in part (a), and ignore the multiplications by the quantity 1/2 in eq. (P5.54–3).]

(c) If, like N, $N/2$ is even, then $f[n]$ and $g[n]$ can each be decomposed into sequences of even- and odd-indexed samples, and therefore, their DFTs can be computed using the same process as in eq. (P5.54–3). Furthermore, if N is an integer power of 2, we can continue to iterate the process, thus achieving significant savings in computation time. With this procedure, approximately how many complex multiplications are required for $N = 32, 256, 1{,}024$, and $4{,}096$? Compare this to the direct method of calculation in part (a).

5.55. In this problem we introduce the concept of *windowing,* which is of great importance both in the design of LTI systems and in the spectral analysis of signals. Windowing is the operation of taking a signal $x[n]$ and multiplying it by a finite-duration *window signal* $w[n]$. That is,

$$p[n] = x[n]w[n].$$

Note that $p[n]$ is also of finite duration.

The importance of windowing in spectral analysis stems from the fact that in numerous applications one wishes to compute the Fourier transform of a signal that has been measured. Since in practice we can measure a signal $x[n]$ only over a finite time interval (the *time window*), the actual signal available for spectral analysis is

$$p[n] = \begin{cases} x[n], & -M \leq n \leq M \\ 0, & \text{otherwise} \end{cases},$$

where $-M \leq n \leq M$ is the time window. Thus,

$$p[n] = x[n]w[n],$$

where $w[n]$ is the *rectangular window;* that is,

$$w[n] = \begin{cases} 1, & -M \leq n \leq M \\ 0, & \text{otherwise} \end{cases}. \tag{P5.55–1}$$

Windowing also plays a role in LTI system design. Specifically, for a variety of reasons (such as the potential utility of the FFT algorithm; see Problem P5.54), it is often advantageous to design a system that has an impulse response of finite duration to achieve some desired signal-processing objective. That is, we often begin with a desired frequency response $H(e^{j\omega})$ whose inverse transform $h[n]$ is an impulse response of infinite (or at least excessively long) duration. What is required then is the construction of an impulse response $g[n]$ of finite duration whose transform $G(e^{j\omega})$ adequately approximates $H(e^{j\omega})$. One general approach to choosing $g[n]$ is to find a window function $w[n]$ such that the transform of $h[n]w[n]$ meets the desired specifications for $G(e^{j\omega})$.

Clearly, the windowing of a signal has an effect on the resulting spectrum. In this problem, we illustrate that effect.

(a) To gain some understanding of the effect of windowing, consider windowing the signal

$$x[n] = \sum_{k=-\infty}^{\infty} \delta[n-k]$$

using the rectangular window signal given in eq. (P5.55–1).

(i) What is $X(e^{j\omega})$?

(ii) Sketch the transform of $p[n] = x[n]w[n]$ when $M = 1$.

(iii) Do the same for $M = 10$.

(b) Next, consider a signal $x[n]$ whose Fourier transform is specified by

$$X(e^{j\omega}) = \begin{cases} 1, & |\omega| < \pi/4 \\ 0, & \pi/4 < |\omega| \le \pi \end{cases}.$$

Let $p[n] = x[n]w[n]$, where $w[n]$ is the rectangular window of eq. (P5.55–1). Roughly sketch $P(e^{j\omega})$ for $M = 4$, 8, and 16.

(c) One of the problems with the use of a rectangular window is that it introduces ripples in the transform $P(e^{j\omega})$. (This is in fact directly related to the Gibbs phenomenon.) For that reason, a variety of other window signals have been developed. These signals are tapered; that is, they go from 0 to 1 more gradually than the abrupt transition of the rectangular window. The result is a reduction in the *amplitude* of the ripples in $P(e^{j\omega})$ at the expense of adding a bit of distortion in terms of further smoothing of $X(e^{j\omega})$.

To illustrate the points just made, consider the signal $x[n]$ described in part (b), and let $p[n] = x[n]w[n]$, where $w[n]$ is the *triangular* or *Bartlett window;* that is,

$$w[n] = \begin{cases} 1 - \frac{|n|}{M+1}, & -M \le n \le M \\ 0, & \text{otherwise} \end{cases}.$$

Roughly sketch the Fourier transform of $p[n] = x[n]w[n]$ for $M = 4$, 8, and 16. [*Hint:* Note that the triangular signal can be obtained as a convolution of a rectangular signal with itself. This fact leads to a convenient expression for $W(e^{j\omega})$.]

(d) Let $p[n] = x[n]w[n]$, where $w[n]$ is a raised cosine signal known as the *Hanning window;* i.e.,

$$w[n] = \begin{cases} \frac{1}{2}[1 + \cos(\pi n/M)], & -M \leq n \leq M \\ 0, & \text{otherwise} \end{cases}.$$

Roughly sketch $P(e^{j\omega})$ for $M = 4, 8$, and 16.

5.56. Let $x[m, n]$ be a signal that is a function of the two independent, discrete variables m and n. In analogy with one dimension and with the continuous-time case treated in Problem 4.53, we can define the two-dimensional Fourier transform of $x[m, n]$ as

$$X(e^{j\omega_1}, e^{j\omega_2}) = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} x[m, n]e^{-j(\omega_1 m + \omega_2 n)}. \qquad \text{(P5.56–1)}$$

(a) Show that eq. (P5.56–1) can be calculated as two successive one-dimensional Fourier transforms, first in m, with n regarded as fixed, and then in n. Use this result to determine an expression for $x[m, n]$ in terms of $X(e^{j\omega_1}, e^{j\omega_2})$.

(b) Suppose that

$$x[m, n] = a[m]b[n],$$

where $a[m]$ and $b[n]$ are each functions of only one independent variable. Let $A(e^{j\omega})$ and $B(e^{j\omega})$ denote the Fourier transforms of $a[m]$ and $b[n]$, respectively. Express $X(e^{j\omega_1}, e^{j\omega_2})$ in terms of $A(e^{j\omega})$ and $B(e^{j\omega})$.

(c) Determine the two-dimensional Fourier transforms of the following signals:

(i) $x[m, n] = \delta[m - 1]\delta[n + 4]$

(ii) $x[m, n] = (\frac{1}{2})^{n-m}u[n - 2]u[-m]$

(iii) $x[m, n] = (\frac{1}{2})^{n}\cos(2\pi m/3)u[n]$

(iv) $x[m, n] = \begin{cases} 1, & -2 < m < 2 \text{ and } -4 < n < 4 \\ 0, & \text{otherwise} \end{cases}$

(v) $x[m, n] = \begin{cases} 1, & -2 + n < m < 2 + n \text{ and } -4 < n < 4 \\ 0, & \text{otherwise} \end{cases}$

(vi) $x[m, n] = \sin\left(\frac{\pi n}{3} + \frac{2\pi m}{5}\right)$

(d) Determine the signal $x[m, n]$ whose Fourier transform is

$$X(e^{j\omega_1}, e^{j\omega_2}) = \begin{cases} 1, & 0 < |\omega_1| \leq \pi/4 \text{ and } 0 < |\omega_2| \leq \pi/2 \\ 0, & \pi/4 < |\omega_1| < \pi \text{ or } \pi/2 < |\omega_2| < \pi \end{cases}.$$

(e) Let $x[m, n]$ and $h[m, n]$ be two signals whose two-dimensional Fourier transforms are denoted by $X(e^{j\omega_1}, e^{j\omega_2})$ and $H(e^{j\omega_1}, e^{j\omega_2})$, respectively. Determine the transforms of the following signals in terms of $X(e^{j\omega_1}, e^{j\omega_2})$ and $H(e^{j\omega_1}, e^{j\omega_2})$:

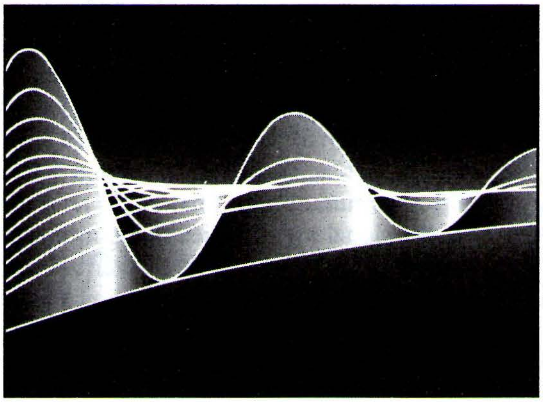
(i) $x[m, n]e^{jW_1 m}e^{jW_2 n}$

(ii) $y[m, n] = \begin{cases} x[k, r], & \text{if } m = 2k \text{ and } n = 2r \\ 0, & \text{if } m \text{ is not a multiple of 2 or } n \text{ is not a multiple of 3} \end{cases}$

(iii) $y[m, n] = x[m, n]h[m, n]$

6

TIME AND FREQUENCY CHARACTERIZATION OF SIGNALS AND SYSTEMS



6.0 INTRODUCTION

The frequency-domain characterization of an LTI system in terms of its frequency response represents an alternative to the time-domain characterization through convolution. In analyzing LTI systems, it is often particularly convenient to utilize the frequency domain because differential and difference equations and convolution operations in the time domain become algebraic operations in the frequency domain. Moreover, concepts such as frequency-selective filtering are readily and simply visualized in the frequency domain. However, in system design, there are typically both time-domain and frequency-domain considerations. For example, as we briefly discussed in Examples 4.18 and 5.12, and as we will illustrate in more detail in this chapter, significant oscillatory behavior in the impulse response of a frequency-selective filter may be undesirable, and consequently, we may wish to sacrifice the level of frequency selectivity in a filter in order to meet the required tolerances on impulse response behavior. Situations such as this are the rule rather than the exception in practice, as in most applications we would like to specify or constrain certain characteristics of a system in both the time domain and the frequency domain, frequently resulting in conflicting requirements. Hence, in system design and analysis, it is important to relate time-domain and frequency-domain characteristics and trade-offs. Introducing these issues and relationships is the primary focus of the chapter.

6.1 THE MAGNITUDE-PHASE REPRESENTATION OF THE FOURIER TRANSFORM

The Fourier transform is in general complex valued and, as we discussed, can be represented in terms of its real and imaginary components or in terms of magnitude and phase.

The magnitude-phase representation of the continuous-time Fourier transform $X(j\omega)$ is

$$X(j\omega) = |X(j\omega)|e^{j\sphericalangle X(j\omega)}. \tag{6.1}$$

Similarly the magnitude-phase representation of the discrete-time Fourier transform $X(e^{j\omega})$ is

$$X(e^{j\omega}) = |X(e^{j\omega})|e^{j\sphericalangle X(e^{j\omega})}. \tag{6.2}$$

In the following discussion, we concentrate for the most part on the continuous-time case in describing and illustrating several points related to magnitude-phase representations. The essential points apply equally to the discrete-time case.

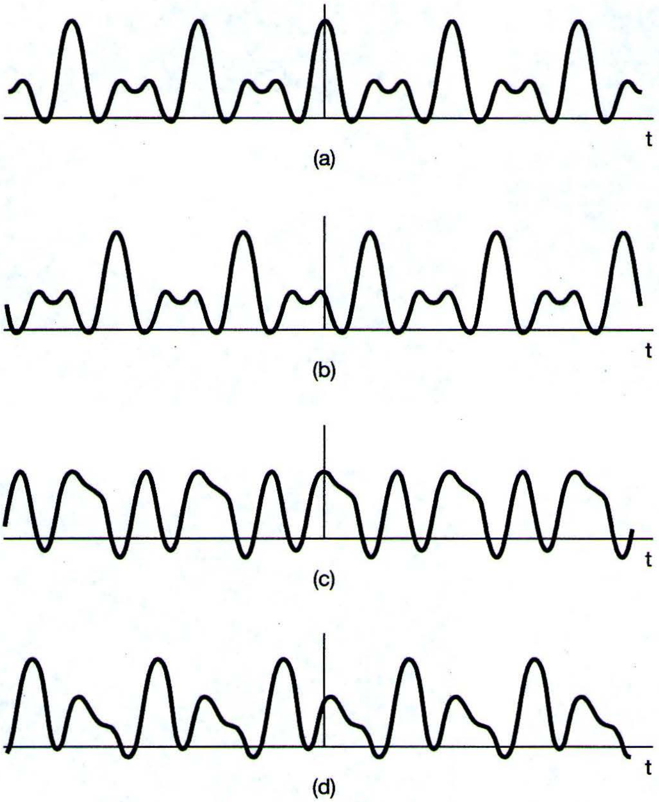
From the Fourier transform synthesis equation (4.8), we can think of $X(j\omega)$ as providing us with a decomposition of the signal $x(t)$ into a "sum" of complex exponentials at different frequencies. In fact, as discussed in Section 4.3.7, $|X(j\omega)|^2$ may be interpreted as the energy-density spectrum of $x(t)$. That is, $|X(j\omega)|^2 d\omega/2\pi$ can be thought of as the amount of energy in the signal $x(t)$ that lies in the infinitesimal frequency band between ω and $\omega + d\omega$. Thus, the magnitude $|X(j\omega)|$ describes the basic frequency content of a signal—i.e., $|X(j\omega)|$ provides us with the information about the relative magnitudes of the complex exponentials that make up $x(t)$. For example, if $|X(j\omega)| = 0$ outside of a small band of frequencies centered at zero, then $x(t)$ will display only relatively low-frequency oscillations.

The phase angle $\sphericalangle X(j\omega)$, on the other hand, does not affect the amplitudes of the individual frequency components, but instead provides us with information concerning the relative phases of these exponentials. The phase relationships captured by $\sphericalangle X(j\omega)$ have a significant effect on the nature of the signal $x(t)$ and thus typically contain a substantial amount of information about the signal. In particular, depending upon what this phase function is, we can obtain very different-looking signals, even if the magnitude function remains unchanged. For example, consider again the example illustrated in Figure 3.3. In this case, a ship encounters the superposition of three wave trains, each of which can be modeled as a sinusoidal signal. With fixed magnitudes for these sinusoids, the amplitude of their sum may be quite small or very large, depending on the relative phases. The implications of phase for the ship, therefore, are quite significant. As another illustration of the effect of phase, consider the signal

$$x(t) = 1 + \frac{1}{2}\cos(2\pi t + \phi_1) + \cos(4\pi t + \phi_2) + \frac{2}{3}\cos(6\pi t + \phi_3). \tag{6.3}$$

In Figure 3.4, we depicted $x(t)$ for the case when $\phi_1 = \phi_2 = \phi_3 = 0$. In Figure 6.1, we illustrate $x(t)$ for this case also and for several other choices for the phase of the individual components. As this figure demonstrates, the resulting signals can differ significantly for different relative phases.

In general, changes in the phase function of $X(j\omega)$ lead to changes in the time-domain characteristics of the signal $x(t)$. In some instances phase distortion may be important, whereas in others it is not. For example, a well-known property of the auditory system is a relative insensitivity to phase. Specifically, if the Fourier transform of a spoken sound (e.g., a vowel) is subjected to a distortion such that the phase is changed but the magnitude is unchanged, the effect can be perceptually negligible, although the waveform in the time domain may look considerably different. While mild phase distortions such as those affecting individual sounds do not lead to a loss of intelligibility, more severe phase



(a)

(b)

(c)

(d)

Figure 6.1 The signal $x(t)$ given in eq. (6.3) for several different choices of the phase angles ϕ_1, ϕ_2, and ϕ_3: (a) $\phi_1 = \phi_2 = \phi_3 = 0$; (b) $\phi_1 = 4$ rad, $\phi_2 = 8$ rad, $\phi_3 = 12$ rad; (c) $\phi_1 = 6$ rad, $\phi_2 = -2.7$ rad, $\phi_3 = 0.93$ rad; (d) $\phi_1 = 1.2$ rad, $\phi_2 = 4.1$ rad, $\phi_3 = -7.02$ rad.

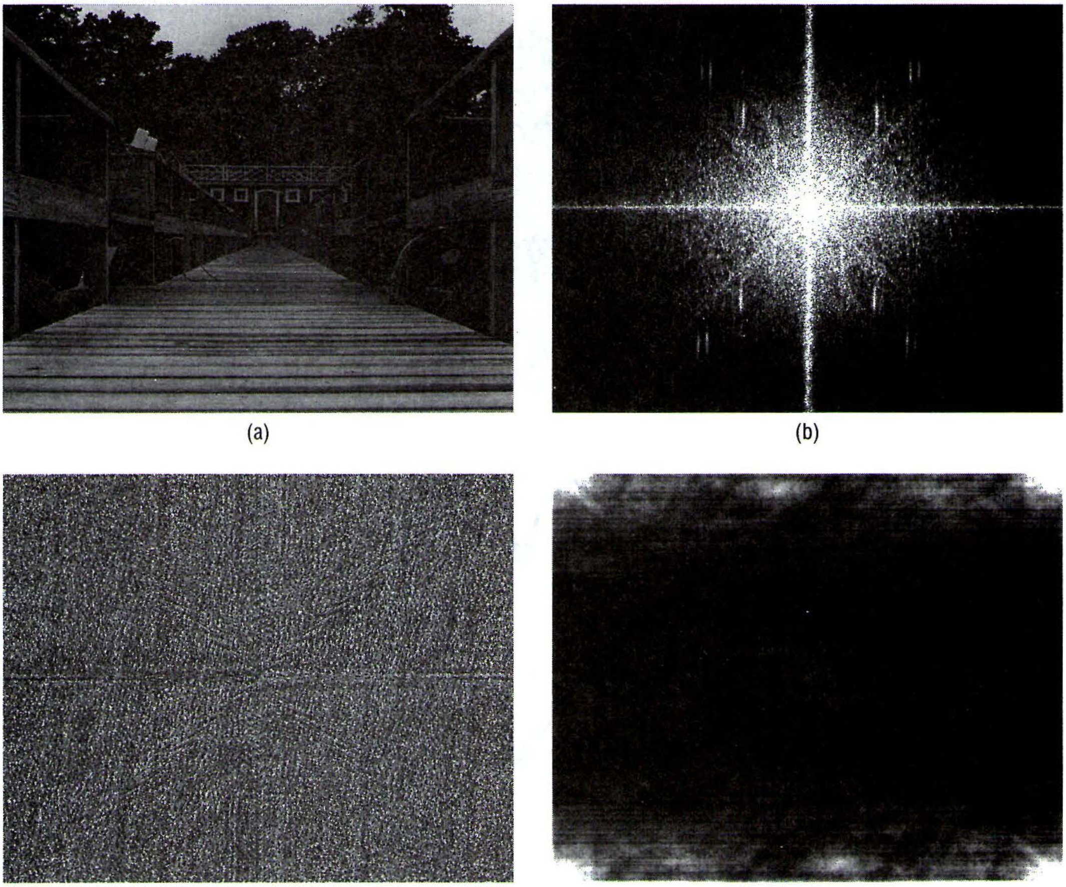
distortions of speech certainly do. As an extreme illustration, if $x(t)$ is a tape recording of a sentence, then the signal $x(-t)$ represents the sentence played backward. From Table 4.1, assuming $x(t)$ is real valued, the corresponding effect in the frequency domain is to replace the Fourier transform phase by its negative:

$$\mathfrak{F}\{x(-t)\} = X(-j\omega) = |X(j\omega)|e^{-j\sphericalangle X(j\omega)}.$$

That is, the spectrum of a sentence played in reverse has the same magnitude function as the spectrum of the original sentence and differs only in phase. Clearly, this phase change has a significant impact on the intelligibility of the recording.

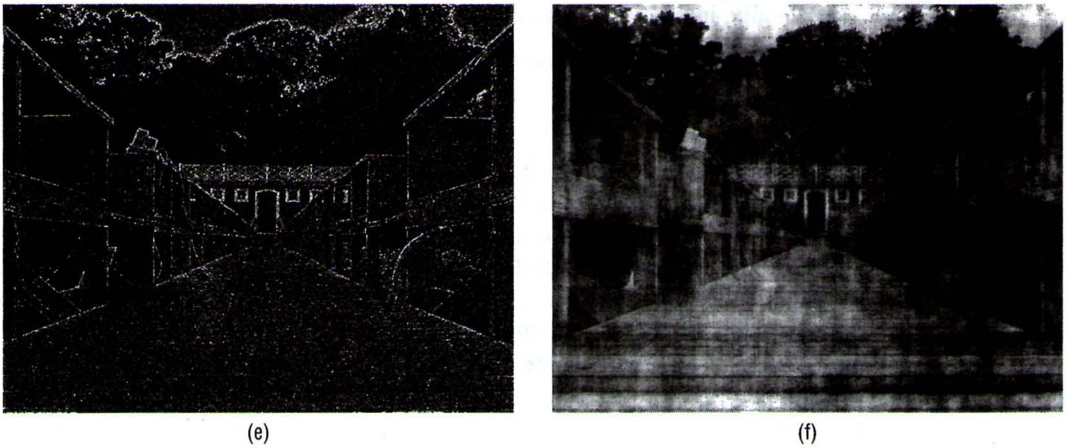
A second example illustrating the effect and importance of phase is found in examining images. As we briefly discussed in Chapter 3, a black-and-white picture can be thought of as a signal $x(t_1, t_2)$, with t_1 denoting the horizontal coordinate of a point on the picture, t_2 the vertical coordinate, and $x(t_1, t_2)$ the brightness of the image at the point (t_1, t_2). The Fourier transform $X(j\omega_1, j\omega_2)$ of the image represents a decomposition of the image into complex exponential components of the form $e^{j\omega_1 t_1} e^{j\omega_2 t_2}$ that capture the spatial variations of $x(t_1, t_2)$ at different frequencies in each of the two coordinate directions. Several elementary aspects of two-dimensional Fourier analysis are addressed in Problems 4.53 and 5.56.

In viewing a picture, some of the most important visual information is contained in the edges and regions of high contrast. Intuitively, regions of maximum and minimum


(a) (b)

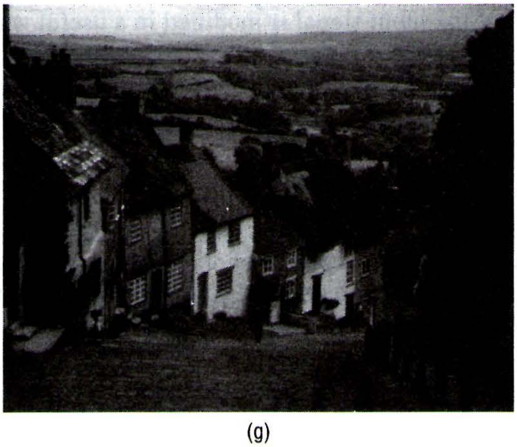
(c) (d)

intensity in a picture are places at which complex exponentials at different frequencies are in phase. Therefore, it seems plausible to expect the phase of the Fourier transform of a picture to contain much of the information in the picture, and in particular, the phase should capture the information about the edges. To substantiate this expectation, in Figure 6.2(a) we have repeated the picture shown in Figure 1.4. In Figure 6.2(b) we have depicted the magnitude of the two-dimensional Fourier transform of the image in Figure 6.2(a), where in this image the horizontal axis is ω_1, the vertical is ω_2, and the brightness of the image at the point (ω_1, ω_2) is proportional to the magnitude of the transform $X(j\omega_1, j\omega_2)$ of the image in Figure 6.2(a). Similarly, the phase of this transform is depicted in Figure 6.2(c). Figure 6.2(d) is the result of setting the phase [Figure 6.2(c)] of $X(j\omega_1, j\omega_2)$ to zero (without changing its magnitude) and inverse transforming. In Figure 6.2(e) the magnitude of $X(j\omega_1, j\omega_2)$ was set equal to 1, but the phase was kept unchanged from what it was in Figure 6.2(c). Finally, in Figure 6.2(f) we have depicted the image obtained by inverse transforming the function obtained by using the phase in Figure 6.2(c) and the magnitude of the transform of a *completely different* image—the picture shown in Figure 6.2(g)! These figures clearly illustrate the importance of phase in representing images.



(e)

(f)



(g)

Figure 6.2 (a) The image shown in Figure 1.4; (b) magnitude of the two-dimensional Fourier transform of (a); (c) phase of the Fourier transform of (a); (d) picture whose Fourier transform has magnitude as in (b) and phase equal to zero; (e) picture whose Fourier transform has magnitude equal to 1 and phase as in (c); (f) picture whose Fourier transform has phase as in (c) and magnitude equal to that of the transform of the picture shown in (g).

6.2 THE MAGNITUDE-PHASE REPRESENTATION OF THE FREQUENCY RESPONSE OF LTI SYSTEMS

From the convolution property for continuous-time Fourier transforms, the transform $Y(j\omega)$ of the output of an LTI system is related to the transform $X(j\omega)$ of the input to the system by the equation

$$Y(j\omega) = H(j\omega)X(j\omega),$$

where $H(j\omega)$ is the frequency response of the system—i.e., the Fourier transform of the system's impulse response. Similarly, in discrete time, the Fourier transforms of the input $X(e^{j\omega})$ and ouput $Y(e^{j\omega})$ of an LTI system with frequency response $H(e^{j\omega})$ are related by

$$Y(e^{j\omega}) = H(e^{j\omega})X(e^{j\omega}). \tag{6.4}$$

Thus, the effect that an LTI system has on the input is to change the complex amplitude of each of the frequency components of the signal. By looking at this effect in terms of the magnitude-phase representation, we can understand the nature of the effect in more detail. Specifically, in continuous time,

$$|Y(j\omega)| = |H(j\omega)||X(j\omega)| \tag{6.5}$$

and

$$\sphericalangle Y(j\omega) = \sphericalangle H(j\omega) + \sphericalangle X(j\omega), \tag{6.6}$$

and exactly analogous relationships hold in the discrete-time case. From eq. (6.5), we see that the effect an LTI system has on the magnitude of the Fourier transform of the signal is to scale it by the magnitude of the frequency response. For this reason, $|H(j\omega)|$ (or $|H(e^{j\omega})|$) is commonly referred to as the *gain* of the system. Also, from eq. (6.6), we see that the phase of the input $\sphericalangle X(j\omega)$ is modified by the LTI system by adding the phase $\sphericalangle H(j\omega)$ to it, and $\sphericalangle H(j\omega)$ is typically referred to as the *phase shift* of the system. The phase shift of the system can change the relative phase relationships among the components of the input, possibly resulting in significant modifications to the time domain characteristics of the input even when the gain of the system is constant for all frequencies. The changes in the magnitude and phase that result from the application of an input to an LTI system may be either desirable, if the input signal is modified in a useful way, or undesirable, if the input is changed in an unwanted manner. In the latter case, the effects in eqs. (6.5) and (6.6) are commonly referred to as magnitude and phase *distortions*. In the following sections, we describe several concepts and tools that allow us to understand these effects a bit more thoroughly.

6.2.1 Linear and Nonlinear Phase

When the phase shift at the frequency ω is a linear function of ω, there is a particularly straightforward interpretation of the effect in the time domain. Consider the continuous-time LTI system with frequency response

$$H(j\omega) = e^{-j\omega t_0}, \tag{6.7}$$

so that the system has unit gain and linear phase—i.e.,

$$|H(j\omega)| = 1, \quad \sphericalangle H(j\omega) = -\omega t_0. \tag{6.8}$$

As shown in Example 4.15, the system with this frequency response characteristic produces an output that is simply a time shift of the input—i.e.,

$$y(t) = x(t - t_0). \tag{6.9}$$

In the discrete-time case, the effect of linear phase is similar to that in the continuous-time case when the slope of the linear phase is an integer. Specifically, from Example 5.11, we know that the LTI system with frequency response $e^{-j\omega n_0}$ with linear phase function $-\omega n_0$ produces an ouput that is a simple shift of the input—i.e., $y[n] = x[n-n_0]$. Thus, a linear phase shift with an *integer* slope corresponds to a shift of $x[n]$ by an integer number of samples. When the phase slope is not an integer, the effect in the time domain is somewhat more complex and is discussed in Chapter 7, Section 7.5. Informally, the effect is a time shift of the envelope of the sequence values, but the values themselves may change.

While linear phase shifts lead to very simple and easily understood and visualized changes in a signal, if an input signal is subjected to a phase shift that is a nonlinear function of ω, then the complex exponential components of the input at different frequencies will be shifted in a manner that results in a change in their relative phases. When these exponentials are superimposed, we obtain a signal that may look considerably different from the input signal. This is illustrated in Figure 6.3 in the continuous-time case.

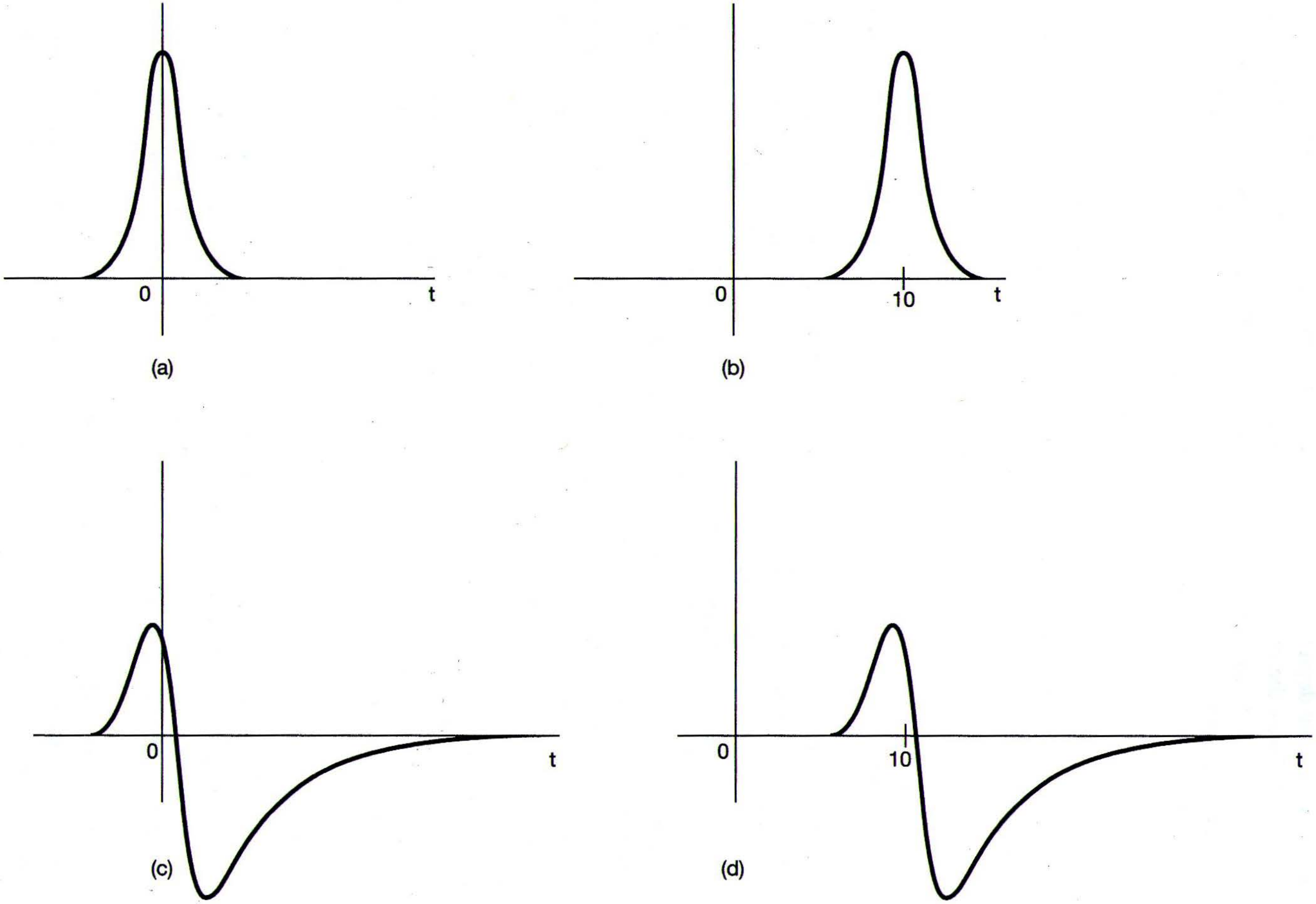


Figure 6.3 (a) Continuous-time signal that is applied as the input to several systems for which the frequency response has unity magnitude; (b) response for a system with linear phase; (c) response for a system with nonlinear phase; and (d) response for a system with phase equal to the nonlinear phase of the system in part (c) plus a linear phase term.

In Figure 6.3(a), we depict a signal that is applied as the input to three different systems. Figure 6.3(b) shows the output when the signal is applied as input to a system with frequency response $H_1(j\omega) = e^{-j\omega t_0}$, resulting in an output that equals the input delayed by t_0 seconds. In Figure 6.3(c), we display the output when the signal is applied to a system with unity gain and nonlinear phase function—i.e.,

$$H_2(j\omega) \quad = \quad e^{j\sphericalangle H_2(j\omega)}, \tag{6.10}$$

where $\sphericalangle H_2(j\omega)$ is a nonlinear function of ω. Figure 6.3(d) shows the output from another system with nonlinear phase. In this case, the corresponding frequency response has a phase shift that is obtained by adding a linear phase term to $\sphericalangle H_2(j\omega)$—i.e.,

$$H_3(j\omega) = H_2(j\omega)e^{-j\omega t_0}. \tag{6.11}$$

Thus, the output in Figure 6.3(d) can be thought of as the response to a cascade of the system $H_2(j\omega)$ followed by a time shift, so that the waveforms in Figures 6.3(c) and (d) are related through a simple time shift.

In Figure 6.4, we illustrate the effect of both linear and nonlinear phase in the discrete-time case. Once again, the signal in Figure 6.4(a) is applied as the input to three different LTI systems, all with unity gain (i.e., $|H(e^{j\omega})| = 1$). The signals in the subsequent parts of Figure 6.4 depict the corresponding outputs. In the case of Figure 6.4(b), the system has linear phase characteristics with integer slope of -5, so that the output equals the input delayed by 5. The phase shifts for the systems associated with Figures 6.4(c) and (d) are nonlinear, but the difference between these two phase functions is linear with integer slope so that the signals in Figures 6.4(c) and (d) are related by a time shift.

Note that all the systems considered in the examples illustrated in Figures 6.3 and 6.4 have unity gain, so that the magnitude of the Fourier transform of the input to any of these systems is passed through *unchanged* by the system. For this reason, such systems are commonly referred to as *all-pass* systems. The characteristics of an all-pass system are completely determined by its phase-shift characteristics. A more general LTI system $H(j\omega)$ or $H(e^{j\omega})$, of course, imparts both magnitude shaping through the gain $|H(j\omega)|$ or $|H(e^{j\omega})|$ and phase shift that may or may not be linear.

6.2.2 Group Delay

As discussed in Section 6.2.1, systems with linear phase characteristics have the particularly simple interpretation as time shifts. In fact, from eqs. (6.8) and (6.9), the phase slope tells us the size of the time shift. That is, in continuous time, if $\sphericalangle H(j\omega) = -\omega t_0$, then the system imparts a time shift of $-t_0$ or, equivalently, a *delay* of t_0. Similarly, in discrete time, $\sphericalangle H(e^{j\omega}) = -\omega n_0$ corresponds to a delay of n_0.

The concept of delay can be very naturally and simply extended to include nonlinear phase characteristics. Suppose that we wish to examine the effects of the phase of a continuous-time LTI system on a narrowband input—i.e., an input $x(t)$ whose Fourier transform is zero or negligibly small outside a small band of frequencies centered at $\omega = \omega_0$. By taking the band to be very small, we can accurately approximate the phase of this system in the band with the linear approximation

$$\sphericalangle H(j\omega) \simeq -\phi - \omega\alpha, \tag{6.12}$$

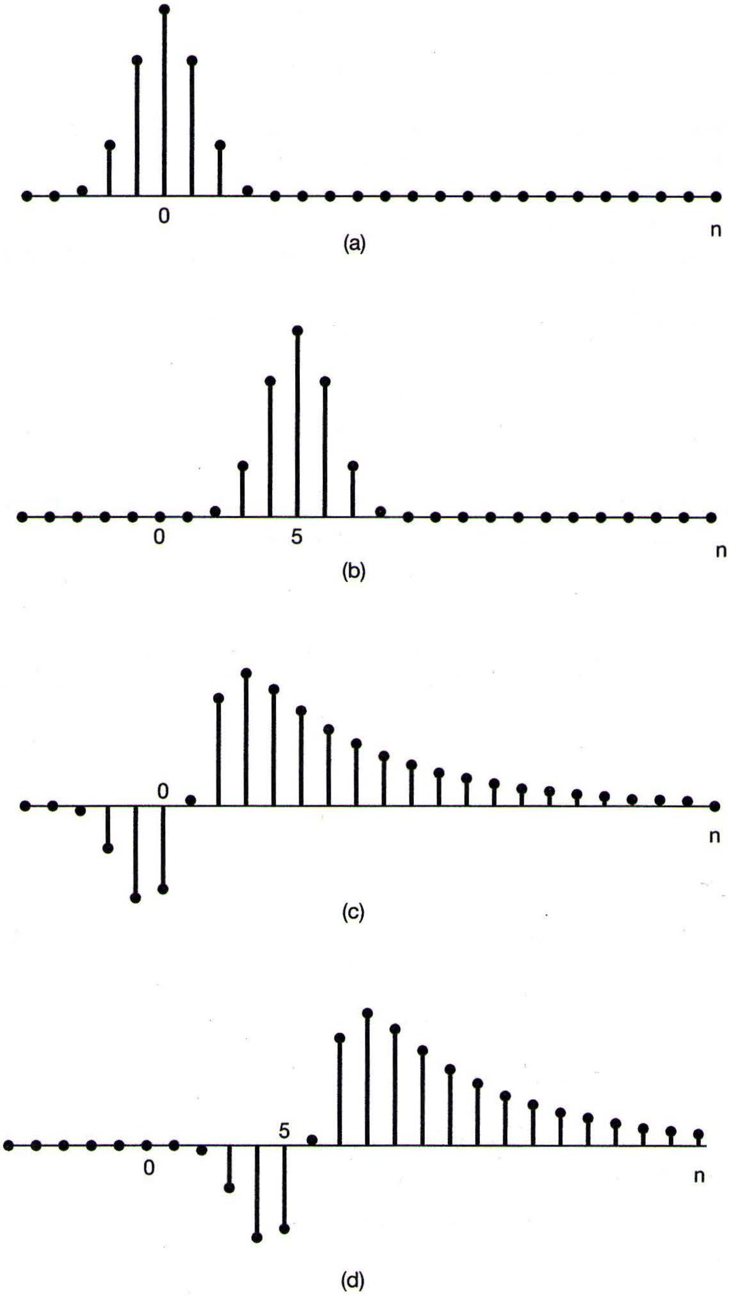


Figure 6.4 (a) Discrete-time signal that is applied as input to several systems for which the frequency response has unity magnitude; (b) response for a system with linear phase with slope of −5; (c) response for a system with nonlinear phase; and (d) response for a system whose phase characteristic is that of part (c) plus a linear phase term with integer slope.

so that

$$Y(j\omega) \simeq X(j\omega)|H(j\omega)|e^{-j\phi}e^{-j\omega\alpha}. \tag{6.13}$$

Thus, the approximate effect of the system on the Fourier transform of this narrowband input consists of the magnitude shaping corresponding to $|H(j\omega)|$, multiplication by an overall constant complex factor $e^{-j\phi}$ and multiplication by a linear phase term $e^{-j\omega\alpha}$ corresponding to a time delay of α seconds. This time delay is referred to as the *group delay* at $\omega = \omega_0$, as it is the effective common delay experienced by the small band or group of frequencies centered at $\omega = \omega_0$.

The group delay at each frequency equals the negative of the slope of the phase at that frequency; i.e., the group delay is defined as

$$\tau(\omega) = -\frac{d}{d\omega}\{\sphericalangle H(j\omega)\}. \tag{6.14}$$

The concept of group delay applies directly to discrete-time systems as well. In the next example we illustrate the effect of nonconstant group delay on a signal.

Example 6.1

Consider the impulse response of an all-pass system with a group delay that varies with frequency. The frequency response $H(j\omega)$ for our example is the product of three factors; i.e.,

$$H(j\omega) = \prod_{i=1}^{3} H_i(j\omega),$$

where

$$H_i(j\omega) = \frac{1 + (j\omega/\omega_i)^2 - 2j\zeta_i(\omega/\omega_i)}{1 + (j\omega/\omega_i)^2 + 2j\zeta_i(\omega/\omega_i)}, \tag{6.15}$$

$$\begin{cases} \omega_1 = 315 \text{ rad/sec and } \zeta_1 = 0.066, \\ \omega_2 = 943 \text{ rad/sec and } \zeta_2 = 0.033, \\ \omega_3 = 1888 \text{ rad/sec and } \zeta_3 = 0.058. \end{cases}$$

It is often useful to express the frequencies ω_i measured in radians per second in terms of frequencies f_i measured in Hertz, where

$$\omega_i = 2\pi f_i.$$

In this case,

$$f_1 \simeq 50 \text{ Hz}$$
$$f_2 \simeq 150 \text{ Hz}$$
$$f_3 \simeq 300 \text{ Hz}.$$

Since the numerator of each of the factors $H_i(j\omega)$ is the complex conjugate of the corresponding denominator, it follows that $|H_i(j\omega)| = 1$. Consequently, we may also conclude that

$$|H(j\omega)| = 1.$$

The phase for each $H_i(j\omega)$ can be determined from eq. (6.15):

$$\sphericalangle H_i(j\omega) = -2\arctan\left[\frac{2\zeta_i\left(\omega/\omega_i\right)}{1-\left(\omega/\omega_i\right)^2}\right],$$

and

$$\sphericalangle H(j\omega) = \sum_{i=1}^{3} \sphericalangle H_i(j\omega).$$

If the values of $\sphericalangle H(j\omega)$ are restricted to lie between $-\pi$ and π, we obtain the *principal-phase* function (i.e., the phase modulo 2π), as shown in Figure 6.5(a) where we have plotted the phase versus frequency measured in Hertz. Note that this function contains discontinuities of size 2π at various frequencies, making the phase function nondifferentiable at those points. However, the addition or subtraction of any integer multiple of 2π to the value of the phase at any frequency leaves the original frequency response unchanged. Thus, by appropriately adding or subtracting such integer multiples of 2π from various portions of the principal phase, we obtain the *unwrapped phase* in Figure 6.5(b). The group delay as a function of frequency may now be computed as

$$\tau(\omega) = -\frac{d}{d\omega}\{\sphericalangle[H(j\omega)]\},$$

where $\sphericalangle[H(j\omega)]$ represents the unwrapped-phase function corresponding to $H(j\omega)$. A plot of $\tau(\omega)$ is shown in Figure 6.5(c). Observe that frequencies in the close vicinity of 50 Hz experience greater delay than frequencies in the vicinity of 150 Hz or 300 Hz. The effect of such nonconstant group delay can also be qualitatively observed in the impulse response (see Figure 6.5(d)) of the LTI system. Recall that $\mathfrak{F}\{\delta(t)\} = 1$. The frequency components of the impulse are all aligned in time in such a way that they combine to form the impulse, which is, of course, highly localized in time. Since the all-pass system has nonconstant group delay, different frequencies in the input are delayed by different amounts. This phenomenon is referred to as *dispersion.* In the current example, the group delay is highest at 50 Hz. Consequently, we would expect the latter parts of the impulse response to oscillate at lower frequencies near 50 Hz. This clearly evident in Figure 6.5(d).

Example 6.2

Nonconstant group delay is among the factors considered important for assessing the transmission performance of switched telecommunications networks. In a survey[1] involving locations all across the continental United States, AT&T/Bell System reported group delay characteristics for various categories of toll calls. Figure 6.6 displays some of the results of this study for two such classes. In particular, what is plotted in each curve in Figure 6.6(a) is the nonconstant portion of the group delay for a specific category of toll calls. That is, for each category, a common constant delay corresponding to

---

[1] "Analog Transmission Performance on the Switched Telecommunications Network," by F. P. Duffy and T. W. Thatcher, Jr., in the *Bell System Technical Journal,* vol. 50, no. 4, April, 1971.

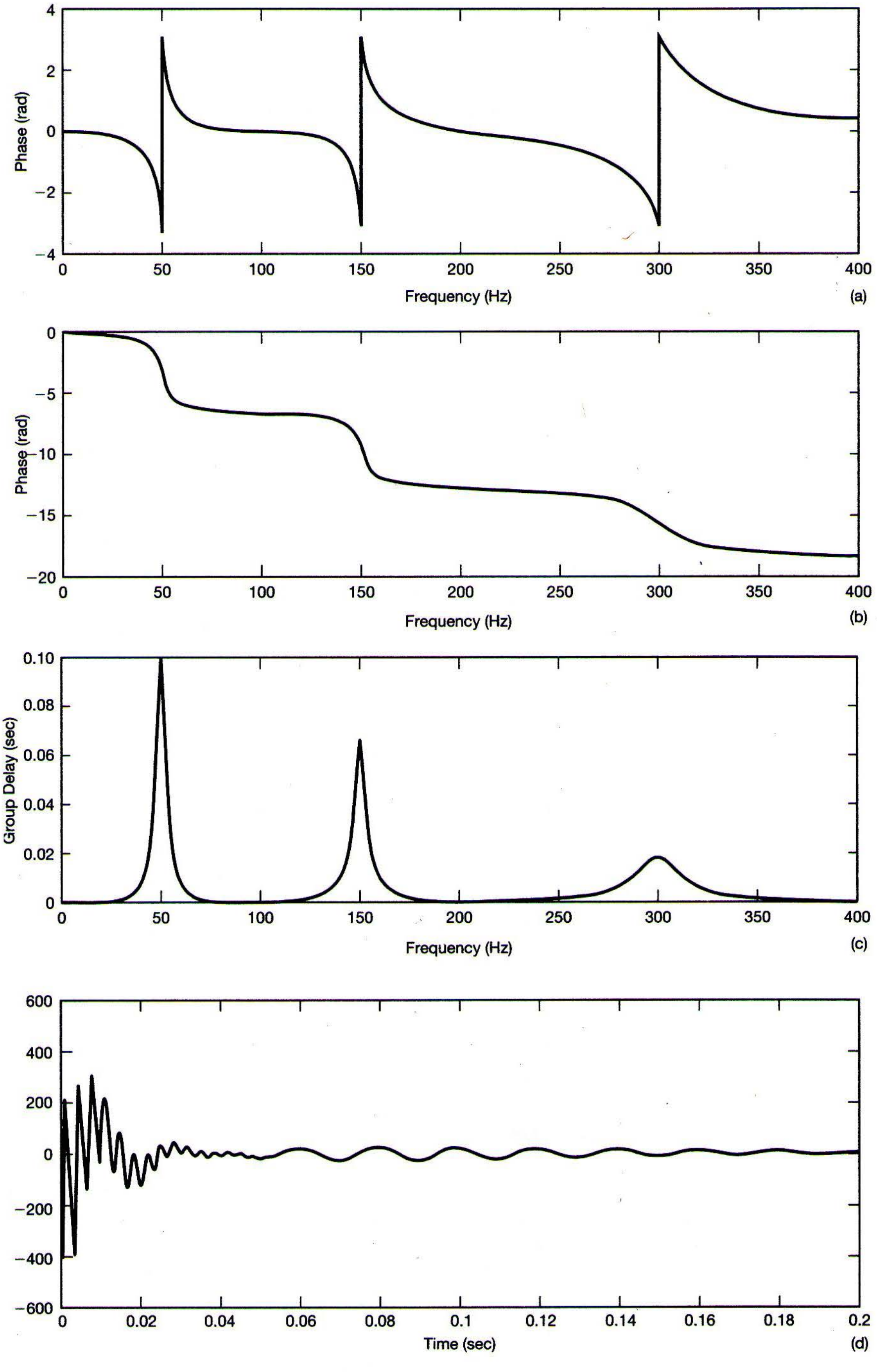


Figure 6.5 Phase, group delay, and impulse response for the all-pass system of Example 6.1: (a) principal phase; (b) unwrapped phase; (c) group delay; (d) impulse response. Each of these quantities is plotted versus frequency measured in Hertz.

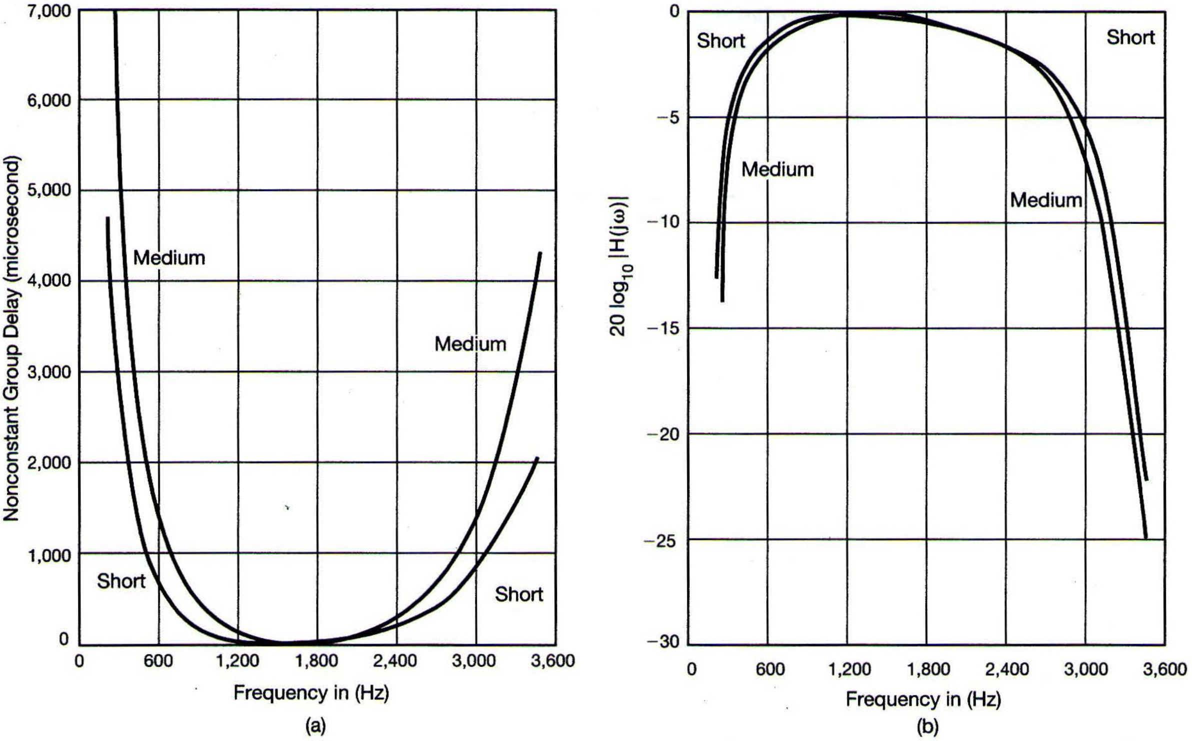
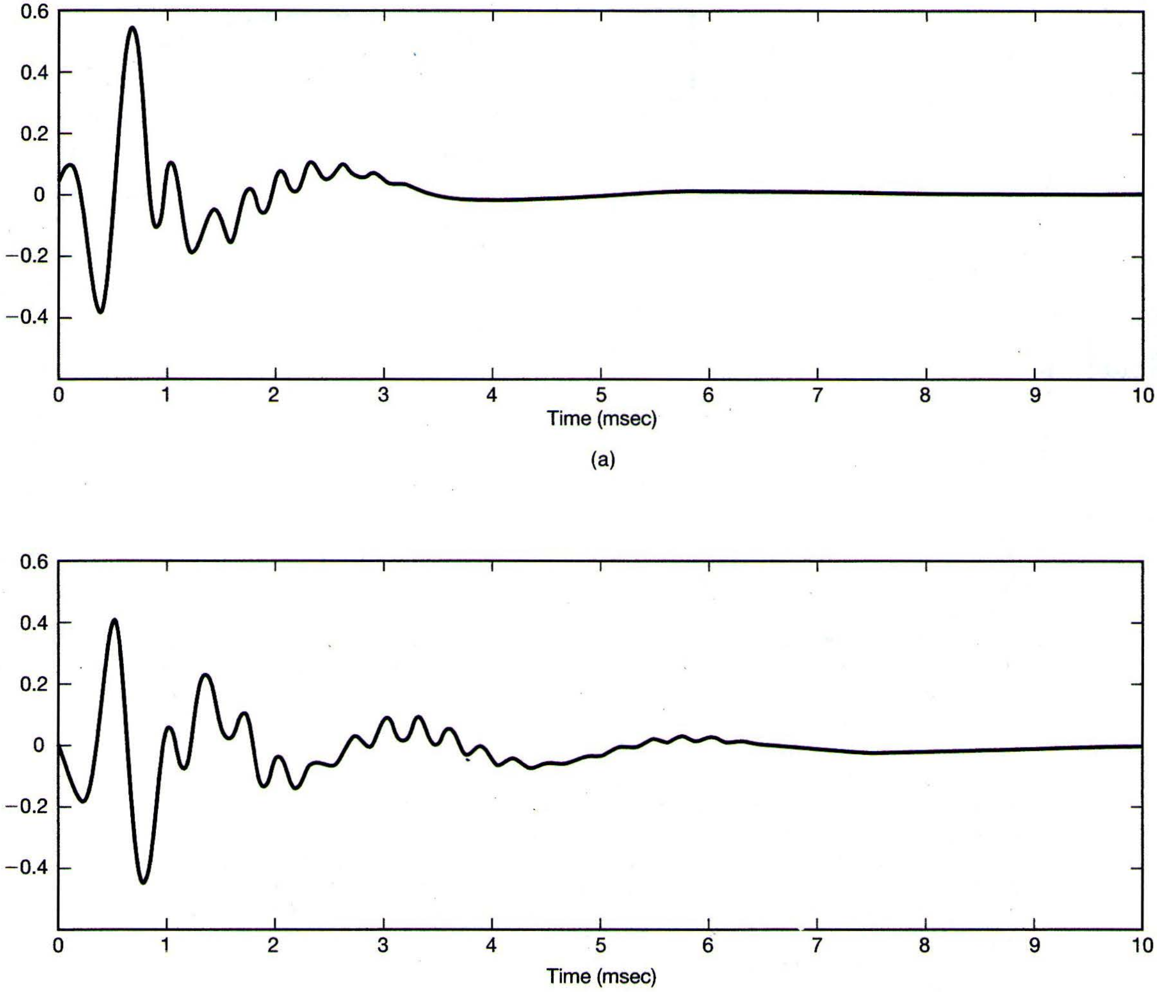


Figure 6.6 (a) Non-constant portion of the group delay; and (b) frequency response magnitude as functions of frequency for short- and medium-distance toll calls in switched telecommunications networks [after Duffy and Thatcher]. Each of these quantities is plotted versus frequency measured in Hertz. Also, as is commonly done in practice, the magnitudes of the frequency responses are plotted using a logarithmic scale in units of *decibels.* That is, what is plotted in (b) is $20 \log_{10} |H(j\omega)|$ for the frequency responses corresponding to short- and medium-distance toll calls. The use of this logarithmic scale for the frequency-response magnitudes is discussed in detail in Section 6.2.3.

the minimum of the group delay over all frequencies has been subtracted from the group delay, and the resulting difference is plotted in Figure 6.6(a). Consequently, each curve in Figure 6.6(a) represents the additional delay (beyond this common constant delay) experienced by the different frequency components of toll calls within each category. The curves labeled SHORT and MEDIUM respectively represent the results for short-distance (0–180 airline miles) and medium-distance (180–725 airline miles) toll calls. The group delay as a function of frequency is seen to be lowest at 1,700 Hz and increases monotonically as we move away from that figure in either direction.

When the group delay characteristics illustrated in Figure 6.6(a) are combined with the characteristics of the magnitude of the frequency response reported in the same AT&T/Bell System survey and shown in Figure 6.6(b), we obtain impulse reponses of the type shown in Figure 6.7. The impulse response in Figure 6.7(a) corresponds to the short-distance category. The very low- and very high-frequency components of the response occur later than the components in the mid-frequency range. This is compatible with



(a)

(b)

Figure 6.7 Impulse responses associated with the group delay and magnitude characteristics in Figure 6.6: (a) impulse response corresponding to the short-distance category of toll calls; (b) impulse response for the medium-distance category.

the corresponding group delay characteristics in Figure 6.6(a). Similarly, Figure 6.7(b) illustrates the same phenomenon for the impulse response corresponding to medium-distance toll calls.

6.2.3 Log-Magnitude and Phase Plots

In graphically displaying continuous-time or discrete-time Fourier transforms and system frequency responses in polar form, it is often convenient to use a logarithmic scale for the magnitude of the Fourier transform. One of the principal reasons for doing this can be seen from eqs. (6.5) and (6.6), which relate the magnitude and phase of the output of an LTI system to those of the input and frequency response. Note that the phase relationship is additive, while the magnitude relationship involves the product of $|H(j\omega)|$ and $|X(j\omega)|$. Thus, if the magnitudes of the Fourier transform are displayed on a logarithmic amplitude scale, eq. (6.5) takes the form of an additive relationship, namely,

$$\log|Y(j\omega)| = \log|H(j\omega)| + \log|X(j\omega)|, \tag{6.16}$$

with an exactly analogous expression in discrete time.

Consequently, if we have a graph of the log magnitude and phase of the Fourier transform of the input and the frequency response of an LTI system, the Fourier transform of the output is obtained by adding the log-magnitude plots and by adding the phase plots. In a similar fashion, since the frequency response of the cascade of LTI systems is the product of the individual frequency responses, we can obtain plots of the log magnitude and phase of the overall frequency response of cascaded systems by adding the corresponding plots for each of the component systems. In addition, plotting the magnitude of the Fourier transform on a logarithmic scale allows detail to be displayed over a wider dynamic range. For example, on a linear scale, the detailed magnitude characteristics in the stopband of a frequency-selective filter with high attenuation are typically not evident, whereas they are on a logarithmic scale.

Typically, the specific logarithmic amplitude scale used is in units of $20\log_{10}$, referred to as *decibels*[2] (abbreviated dB). Thus, 0 dB corresponds to a frequency response with magnitude equal to 1, 20 dB is equivalent to a gain of 10, -20 dB corresponds to an attenuation of 0.1, and so on. Also, it is useful to note that 6 dB approximately corresponds to a gain of 2.

For continuous-time systems, it is also common and useful to use a logarithmic frequency scale. Plots of $20\log_{10}|H(j\omega)|$ and $\sphericalangle H(j\omega)$ versus $\log_{10}(\omega)$ are referred to as *Bode plots.* A typical Bode plot is illustrated in Figure 6.8. Note that, as discussed in Section 4.3.3, if $h(t)$ is real, then $|H(j\omega)|$ is an even function of ω and $\sphericalangle H(j\omega)$ is an odd function of ω. Because of this, the plots for negative ω are superfluous and can be obtained immediately from the plots for positive ω. This, of course, makes it possible to plot frequency response characteristics versus $\log_{10}(\omega)$ for $\omega > 0$, as in the figure.

The use of a logarithmic frequency scale offers a number of advantages in continuous time. For example, it often allows a much wider range of frequencies to be displayed than does a linear frequency scale. In addition, on a logarithmic frequency scale, the shape

---

[2] The origin of this particular choice of units and the term *decibels* can be traced to the definition of power ratios in systems. Specifically, since the square of the magnitude of the Fourier transform of a signal can be interpreted as the energy per unit frequency, or power, in a signal, the square of the magnitude, $|H(j\omega)|^2$ or $|H(e^{j\omega})|^2$, of the frequency response of a system can be thought of as the power ratio between the input and the output of an LTI system. In honor of Alexander Graham Bell, the inventor of the telephone, the term *bel* was introduced to indicate a factor of 10 in a power ratio, and *decibel* was used to denote one-tenth of this factor on a logarithmic scale (so that the cascade of 10 systems with 1-dB power ratios each would result in 1 bel of power amplification). Thus, $10\log_{10}|H(j\omega)|^2$ is the number of decibels of power amplification for the frequency response $H(j\omega)$, and this in turn equals $20\log_{10}|H(j\omega)|$ in magnitude amplification.

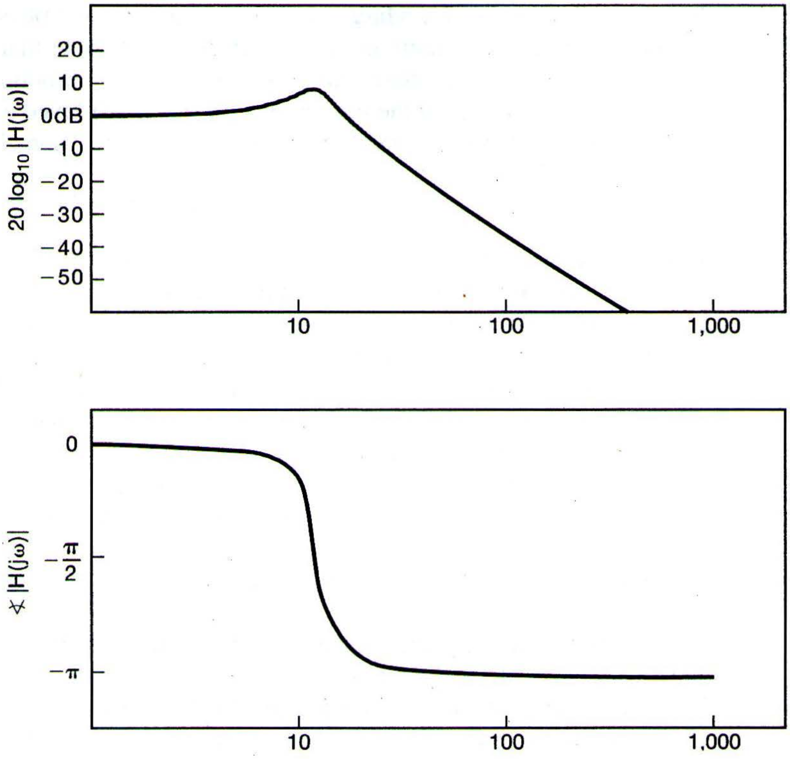


Figure 6.8 A typical Bode plot. (Note that ω is plotted using a logarithmic scale.)

of a particular response curve doesn't change if the frequency is scaled. (See Problem 6.30.) Furthermore for continuous-time LTI systems described by differential equations, an approximate sketch of the log magnitude vs. log frequency can often be easily obtained through the use of asymptotes. In Section 6.5, we will illustrate this by developing simple piecewise-linear approximate Bode plots for first- and second-order continuous-time systems.

In discrete time, the magnitudes of Fourier transforms and frequency responses are often displayed in dB for the same reasons that they are in continuous time. However, in discrete time a logarithmic frequency scale is not typically used, since the range of frequencies to be considered is always limited and the advantage found for differential equations (i.e., linear asymptotes) does not apply to difference equations. Typical graphical representations of the magnitude and phase of a discrete-time frequency response are shown in Figure 6.9. Here, we have plotted $\sphericalangle H(e^{j\omega})$ in radians and $|H(e^{j\omega})|$ in decibels [i.e., $20 \log_{10} |H(e^{j\omega})|$] as functions of ω. Note that for $h[n]$ real, we actually need plot $H(e^{j\omega})$ only for $0 \leq \omega \leq \pi$, because in this case the symmetry property of the Fourier transform implies that we can then calculate $H(e^{j\omega})$ for $-\pi \leq \omega \leq 0$ using the relations $|H(e^{j\omega})| = |H(e^{-j\omega})|$ and $\sphericalangle H(e^{-j\omega}) = -\sphericalangle H(e^{j\omega})$. Furthermore, we need not consider values of $|\omega|$ greater than π, because of the periodicity of $H(e^{j\omega})$.

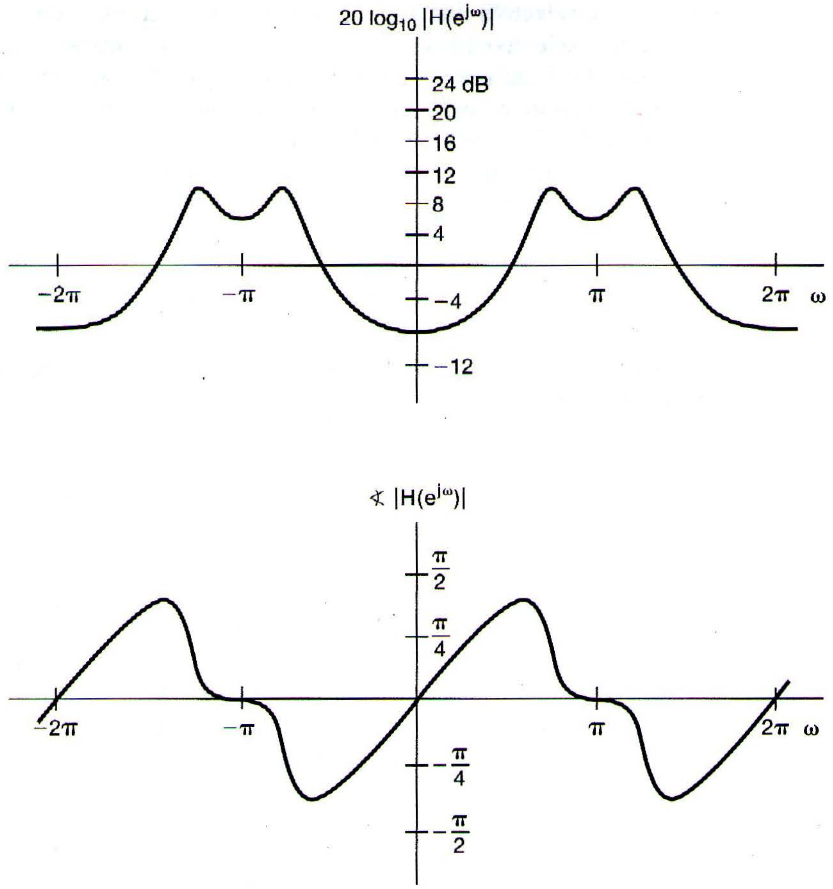


Figure 6.9 Typical graphical representations of the magnitude and phase of a discrete-time frequency response $H(e^{j\omega})$.

As emphasized in this section, a logarithmic amplitude scale is often useful and important. However, there are many situations in which it is convenient to use a linear amplitude scale. For example, in discussing ideal filters for which the magnitude of the frequency response is a nonzero constant over some frequency bands and zero over others, a linear amplitude scale is more appropriate. Thus, we have introduced both linear and logarithmic graphical representations for the magnitude of the Fourier transform and will use each as appropriate.

6.3 TIME-DOMAIN PROPERTIES OF IDEAL FREQUENCY-SELECTIVE FILTERS

In Chapter 3, we introduced the class of frequency-selective filters, i.e., LTI systems with frequency responses chosen so as to pass one or several bands of frequencies with little or no attenuation and to stop or significantly attenuate frequencies outside those bands. As we discussed in Chapters 3, 4, and 5, there are a number of issues of importance that arise in frequency-selective filtering applications and that relate directly to the characteristics of frequency-selective filters. In this section, we take another look at such filters and their properties. We focus our attention here on lowpass filters, although very similar concepts and results hold for other types of frequency-selective filters such as highpass or bandpass filters. (See Problems 6.5, 6.6, 6.26, and 6.38.)

As introduced in Chapter 3, a continuous-time ideal lowpass filter has a frequency response of the form

$$H(j\omega) = \begin{cases} 1 & |\omega| \leq \omega_c \\ 0 & |\omega| > \omega_c \end{cases}. \tag{6.17}$$

This is illustrated in Figure 6.10(a). Similarly, a discrete-time ideal lowpass filter has a frequency response

$$H(e^{j\omega}) = \begin{cases} 1 & |\omega| \leq \omega_c \\ 0 & \omega_c < |\omega| \leq \pi \end{cases} \tag{6.18}$$

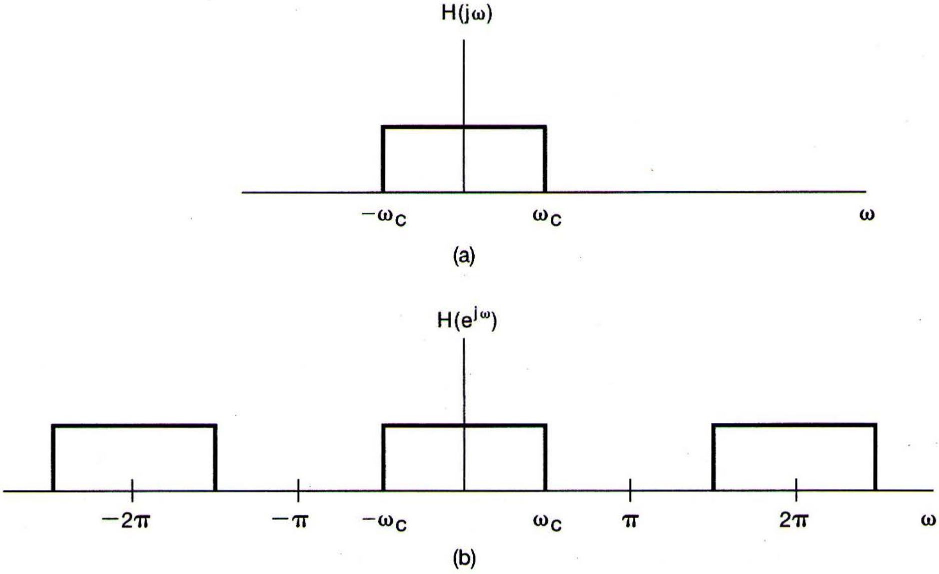


Figure 6.10 (a) The frequency response of a continuous-time ideal lowpass filter; (b) the frequency response of a discrete-time ideal lowpass filter.

and is periodic in ω, as depicted in Figure 6.10(b). As can be seen from eqs. (6.17) and (6.18) or from Figure 6.10, ideal lowpass filters have perfect frequency selectivity. That is, they pass without attenuation all frequencies at or lower than the cutoff frequency ω_c and completely stop all frequencies in the stopband (i.e., higher than ω_c). Moreover, these filters have zero phase characteristics, so they introduce no phase distortion.

As we have seen in Section 6.2, nonlinear phase characteristics can lead to significant changes in the time-domain characteristics of a signal even when the magnitude of its spectrum is not changed by the system, and thus, a filter with a *magnitude* characteristic as in eq. (6.17) or eq. (6.18), but with nonlinear phase, might produce undesirable effects in some applications. On the other hand, an ideal filter with linear phase over the passband, as illustrated in Figure 6.11, introduces only a simple time shift relative to the response of the ideal lowpass filter with zero phase characteristic.

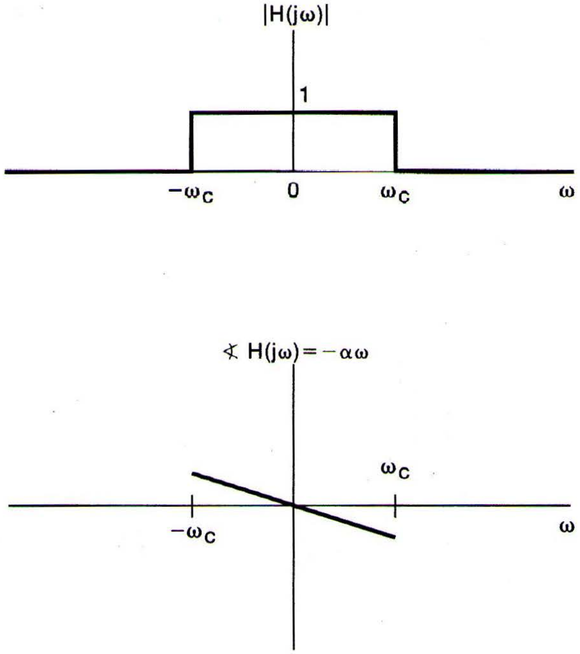


Figure 6.11 Continuous-time ideal lowpass filter with linear phase characteristic.

In Examples 4.18 and 5.12, we computed the impulse responses of ideal lowpass filters. In particular, the impulse response corresponding to the filter in eq. (6.17) is

$$h(t) = \frac{\sin \omega_c t}{\pi t}, \tag{6.19}$$

which is shown in Figure 6.12(a). Similarly, the impulse response of the discrete-time ideal filter in eq. (6.18) is

$$h[n] = \frac{\sin \omega_c n}{\pi n}, \tag{6.20}$$

which is depicted in Figure 6.12(b) for $\omega_c = \pi/4$. If either of the ideal frequency responses of eqs. (6.17) and (6.18) is augmented with a linear phase characteristic, the impulse response is simply delayed by an amount equal to the negative of the slope of this phase function, as is illustrated in Figure 6.13 for the continuous-time impulse response. Note that in both continuous and discrete time, the width of the filter passband is proportional to ω_c, while the width of the main lobe of the impulse is proportional to $1/\omega_c$. As the bandwidth of the filter increases, the impulse response becomes narrower, and vice versa, consistent with the inverse relationship between time and frequency discussed in Chapters 4 and 5.

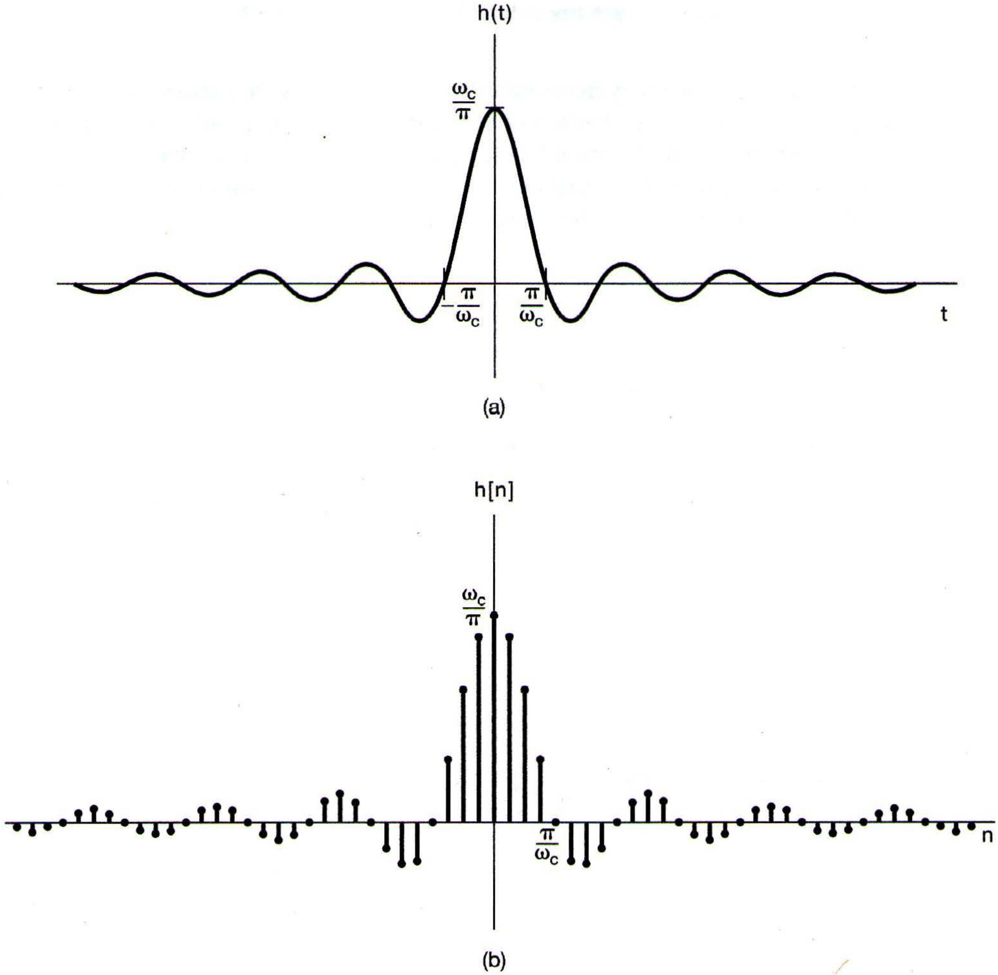


Figure 6.12 (a) The impulse response of the continuous-time ideal lowpass filter of Figure 6.10(a); (b) the impulse response of the discrete-time ideal lowpass filter of Figure 6.10(b) with $\omega_c = \pi/4$.

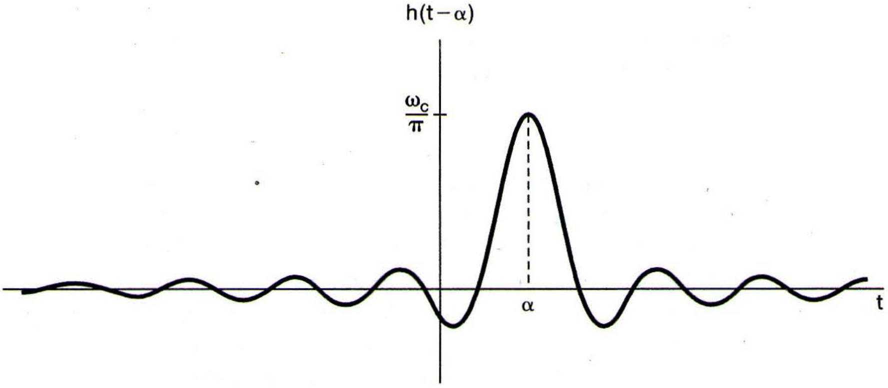


Figure 6.13 Impulse response of an ideal lowpass filter with magnitude and phase shown in Figure 6.11.

The step responses $s(t)$ and $s[n]$ of the ideal lowpass filters in continuous time and discrete time are displayed in Figure 6.14. In both cases, we note that the step responses exhibit several characteristics that may not be desirable. In particular, for these filters, the step responses overshoot their long-term final values and exhibit oscillatory behavior, frequently referred to as *ringing*. Also, recall that the step response is the running integral or sum of the impulse response—i.e.,

$$s(t) = \int_{-\infty}^{t} h(\tau)d\tau,$$

$$s[n] = \sum_{m=-\infty}^{n} h[m].$$

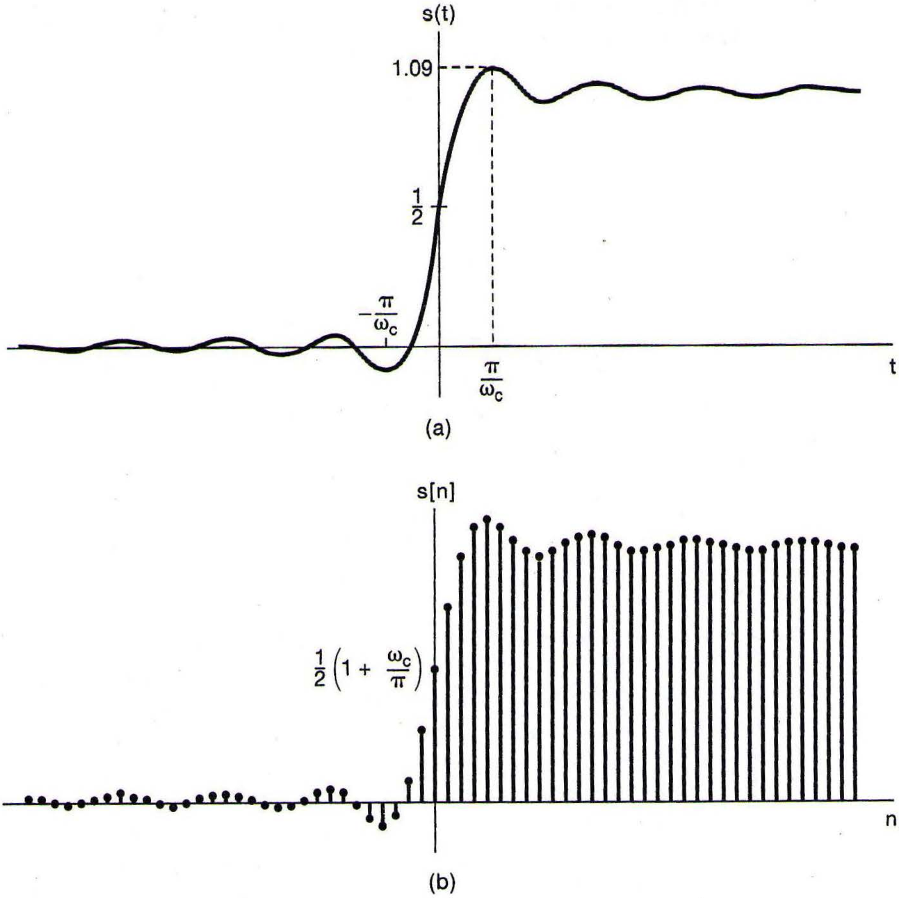


Figure 6.14 (a) Step response of a continuous-time ideal lowpass filter; (b) step response of a discrete-time ideal lowpass filter.

Since the impulse responses for the ideal filters have main lobes extending from $-\pi/\omega_c$ to $+\pi/\omega_c$, the step responses undergo their most significant change in value over this time interval. That is, the so-called *rise time* of the step response, a rough measure of the response time of the filter, is also inversely related to the bandwidth of the filter.

6.4 TIME-DOMAIN AND FREQUENCY-DOMAIN ASPECTS OF NONIDEAL FILTERS

The characteristics of ideal filters are not always desirable in practice. For example, in many filtering contexts, the signals to be separated do not always lie in totally disjoint frequency bands. A typical situation might be that depicted in Figure 6.15, where the spectra of two signals overlap slightly. In such a case, we may wish to trade off the fidelity with which the filter preserves one of these signals—say, $x_1(t)$—against the level to which frequency components of the second signal $x_2(t)$ are attenuated. A filter with a gradual transition from passband to stopband is generally preferable when filtering the superposition of signals with overlapping spectra.

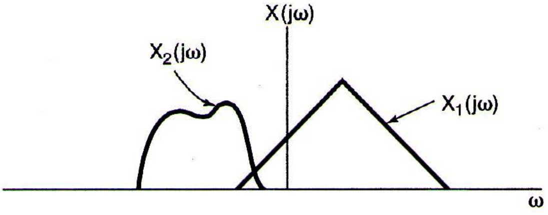


Figure 6.15 Two spectra that are slightly overlapping.

Another consideration is suggested by examining the step responses of ideal lowpass filters, shown in Figure 6.14. For both continuous time and discrete time, the step response asymptotically approaches a constant equal to the value of the step. In the vicinity of the discontinuity, however, it overshoots this value and exhibits ringing. In some situations, this time-domain behavior may be undesirable.

Moreover, even in cases where the ideal frequency-selective characteristics are desirable, they may not be attainable. For example, from eqs. (6.18) and (6.19) and Figure 6.12, it is evident that the ideal lowpass filter is noncausal. When filtering is to be carried out in real time, however, causality is a necessary constraint, and thus, a causal approximation to the ideal characteristics would be required. A further consideration that motivates providing some flexibility in the filter characteristics is ease of implementation. In general, the more precisely we try to approximate or implement an ideal frequency-selective filter, the more complicated or costly the implementation becomes, whether in terms of components such as resistors, capacitors, and operational amplifiers in continuous time or in terms of memory registers, multipliers, and adders in discrete time. In many contexts, a precise filter characteristic may not be essential and a simple filter will suffice.

For all of these reasons, nonideal filters are of of considerable practical importance, and the characteristics of such filters are frequently specified or quantified in terms of several parameters in both the frequency and time domain. First, because the magnitude characteristics of the ideal frequency-selective filter may be unachievable or undesirable, it is preferable to allow some flexibility in the behavior of the filter in the passband and in the stopband, as well as to permit a more gradual transition between the passband and stopband, as opposed to the abrupt transition characteristic of ideal filters. For example, in the case of lowpass filters, the specifications may allow some deviation from unity gain in the passband and from zero gain in the stopband, as well as including both a passband edge and stopband edge with a transition band between them. Thus, specifications for a continuous-time lowpass filter are often stated to require the magnitude of the frequency response of the filter to be restricted to the nonshaded area indicated in Figure 6.16. In this figure, a deviation from unity of plus and minus δ_1 is allowed in the passband, and a deviation of δ_2 from zero is allowed in the stopband. The amount by which the frequency response differs from unity in the passband is referred to as the *passband ripple*, and the amount by which it deviates from zero in the stopband is referred to as the *stopband ripple*. The frequency ω_p is referred to as the *passband edge* and ω_s as the *stopband edge*. The frequency range from ω_p to ω_s is provided for the transition from passband to stopband and is referred to as the *transition band*. Similar definitions apply to discrete-time lowpass filters, as well as to other continuous- and discrete-time frequency-selective filters.

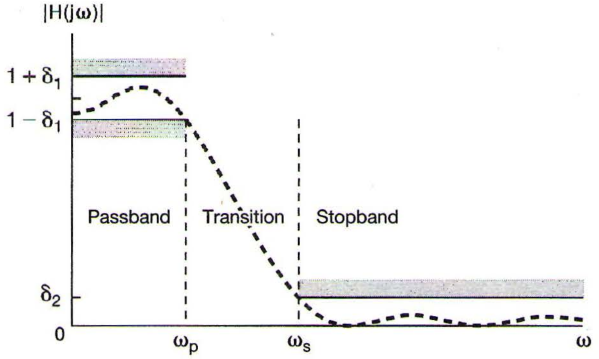


Figure 6.16 Tolerances for the magnitude characteristic of a lowpass filter. The allowable passband ripple is δ_1 and stopband ripple is δ_2. The dashed curve illustrates one possible frequency response that stays within the tolerable limits.

In addition to the specification of magnitude characteristics in the frequency domain, in some cases the specification of phase characteristics is also important. In particular, a linear or nearly linear phase characteristic over the passband of the filter is frequently desirable.

To control the time-domain behavior, specifications are frequently imposed on the step response of a filter. As illustrated in Figure 6.17, one quantity often of interest is the rise time t_r of the step response—i.e., the interval over which the step response rises toward its final value. In addition, the presence or absence of oscillatory behavior, or ringing, in the step response is often of importance. If such ringing is present, then there are three other quantities that are often used to characterize the nature of these oscillations: the overshoot Δ of the final value of the step response, the ringing frequency ω_r, and the settling time t_s—i.e., the time required for the step response to settle to within a specified tolerance of its final value.

For nonideal lowpass filters, a trade-off may be observed between the width of the transition band (a frequency-domain characteristic) and the settling time of the step response (a time-domain characteristic). The following example illustrates this trade-off.

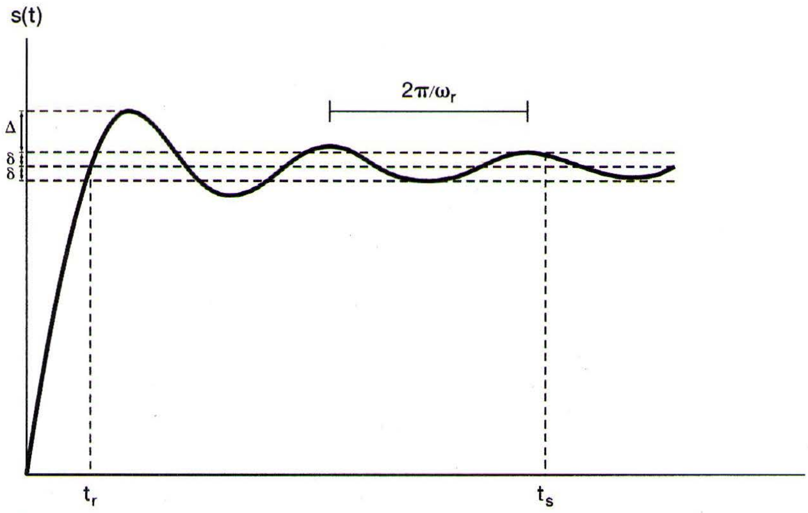


Figure 6.17 Step response of a continuous-time lowpass filter, indicating the rise time t_r, overshoot Δ, ringing frequency ω_r, and settling time t_s—i.e., the time at which the step response settles to within $\pm\delta$ of its final value.

Example 6.3

Let us consider two specific lowpass filters designed to have a cutoff frequency of 500 Hz. Each filter has a fifth-order rational frequency response and a real-valued impulse response. The two filters are of specific types, one referred to as Butterworth filters and the other as elliptic filters. Both of these classes of filters are frequently used in practice.

The magnitudes of the frequency responses of the two filters are plotted (versus frequency measured in Hertz) in Figure 6.18(a). We take the transition band of each filter as the region around the cutoff frequency (500 Hz) where the frequency response magnitude is neither within .05 of unity magnitude (the passband ripple) nor within .05 of zero magnitude (the stopband ripple). From Figure 6.18(a), it can be seen that the transition band of the Butterworth filter is *wider* than the transition band of the elliptic filter.

The price paid for the narrower transition band of the elliptic filter may be observed in Figure 6.18(b), in which the step responses of both filters are displayed. We see that the ringing in the elliptic filter's step response is more prominent than for the Butterworth step response. In particular, the settling time for the step response is longer in the case of the elliptic filter.

The consideration of the trade-offs between time-domain and frequency-domain characteristics and of other issues such as the complexity and cost of filters forms the core of the important field of filter design. In the next few sections, and in several of the problems at the end of the chapter, we provide additional examples of LTI systems and filters and their time- and frequency-domain characteristics.

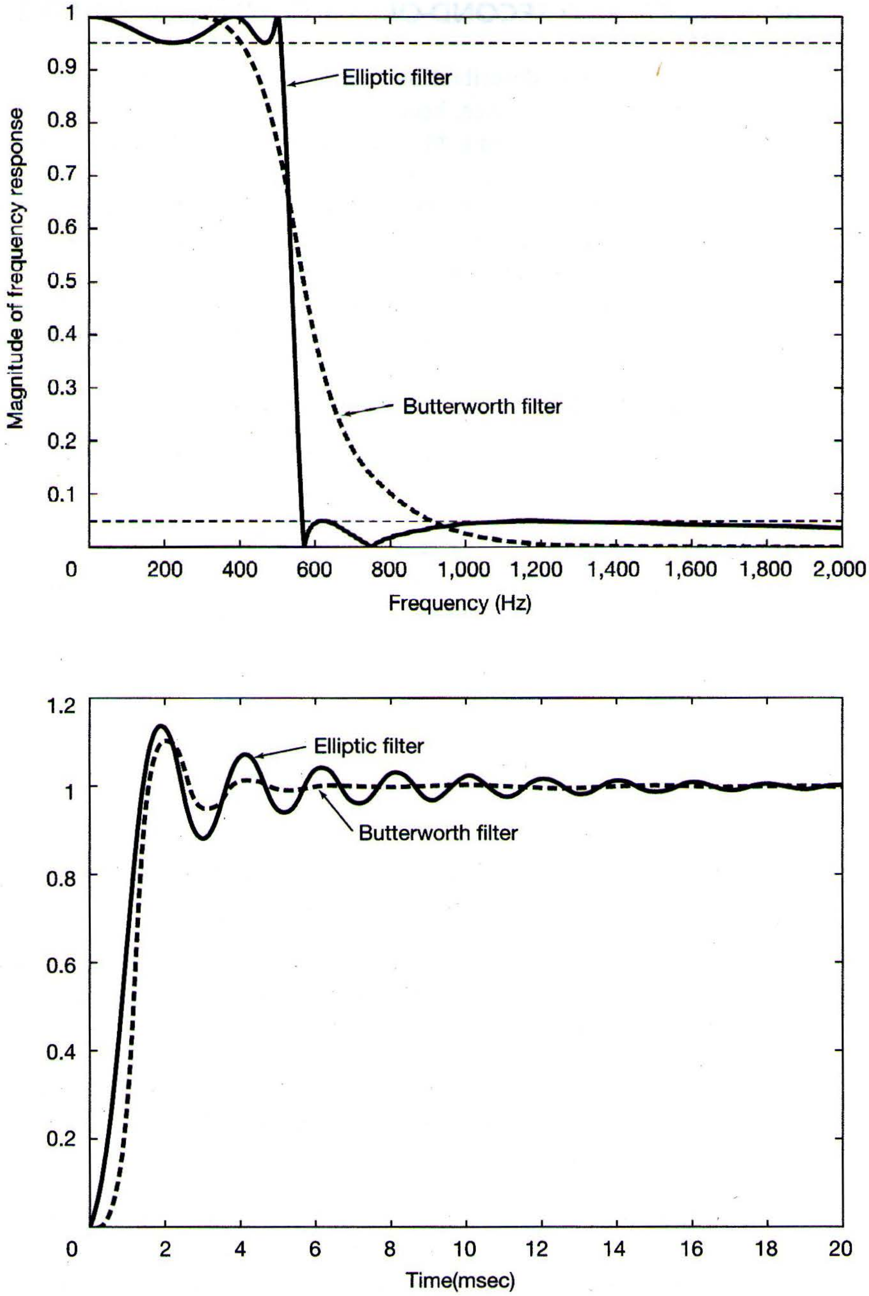


Figure 6.18 Example of a fifth-order Butterworth filter and a fifth-order elliptic filter designed to have the same passband and stopband ripple and the same cutoff frequency: (a) magnitudes of the frequency responses plotted versus frequency measured in Hertz; (b) step responses.

6.5 FIRST-ORDER AND SECOND-ORDER CONTINUOUS-TIME SYSTEMS

LTI systems described by linear constant-coefficient differential equations are of great practical importance, because many physical systems can be modeled by such equations and because systems of this type can often be conveniently implemented. For a variety of practical reasons, high-order systems are frequently implemented or represented by combining first-order and second-order systems in cascade or parallel arrangements. Consequently, the properties of first- and second-order systems play an important role in analyzing, designing, and understanding the time-domain and frequency-domain behavior of higher order systems. In this section, we discuss these low-order systems in detail for continuous time. In Section 6.6, we examine their discrete-time counterparts.

6.5.1 First-Order Continuous-Time Systems

The differential equation for a first-order system is often expressed in the form

$$\tau\frac{dy(t)}{dt} + y(t) = x(t), \tag{6.21}$$

where the coefficient τ is a positive number whose significance will be made clear shortly. The corresponding frequency response for the first-order system is

$$H(j\omega) = \frac{1}{j\omega\tau + 1}, \tag{6.22}$$

and the impulse response is

$$h(t) = \frac{1}{\tau}e^{-t/\tau}u(t), \tag{6.23}$$

which is sketched in Figure 6.19(a). The step response of the system is

$$s(t) = h(t) * u(t) = [1 - e^{-t/\tau}]u(t). \tag{6.24}$$

This is sketched in Figure 6.19(b). The parameter τ is the *time constant* of the system, and it controls the rate at which the first-order system responds. For example, as illustrated in Figure 6.19, at $t = \tau$ the impulse response has reached $1/e$ times its value at $t = 0$, and the step response is within $1/e$ of its final value. Therefore, as τ is decreased, the impulse response decays more sharply, and the rise time of the step response becomes shorter—i.e., it rises more sharply toward its final value. Note also that the step response of a first-order system does not exhibit any ringing.

Figure 6.20 depicts the Bode plot of the frequency response of eq. (6.22). In this figure we illustrate one of the advantages of using a logarithmic frequency scale: We can, without too much difficulty, obtain a useful approximate Bode plot for a continuous-time first-order system. To see this, let us first examine the plot of the log magnitude of the frequency response. Specifically, from eq. (6.22), we obtain

$$20\log_{10}|H(j\omega)| = -10\log_{10}[(\omega\tau)^2 + 1]. \tag{6.25}$$

From this, we see that for $\omega\tau \ll 1$, the log magnitude is approximately zero, while for $\omega\tau \gg 1$, the log magnitude is approximately a *linear* function of $\log_{10}(\omega)$. That is,

$$20\log_{10}|H(j\omega)| \simeq 0 \qquad \text{for} \quad \omega \ll 1/\tau, \tag{6.26}$$

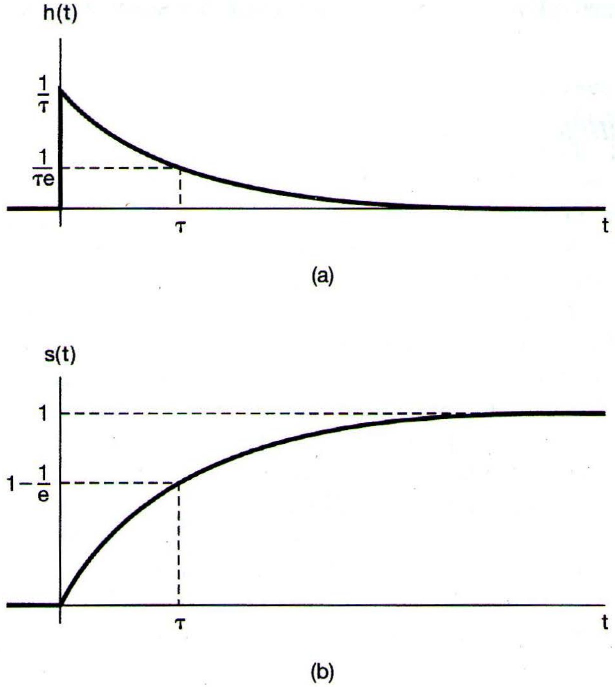


Figure 6.19 Continuous-time first-order system: (a) impulse response; (b) step response.

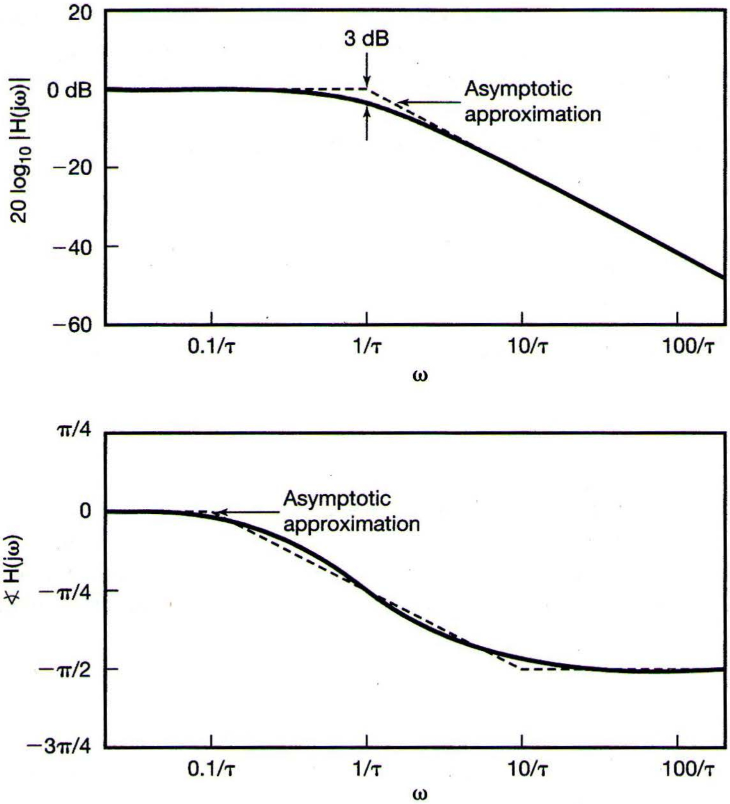


Figure 6.20 Bode plot for a continuous-time first-order system.

and

$$\begin{aligned} 20\log_{10}|H(j\omega)| &\simeq -20\log_{10}(\omega\tau) \\ &= -20\log_{10}(\omega) - 20\log_{10}(\tau) \quad \text{for} \quad \omega \gg 1/\tau. \end{aligned} \tag{6.27}$$

In other words, for the first-order system, the low- and high-frequency asymptotes of the log magnitude are straight lines. The low-frequency asymptote [given by eq. (6.26)] is just the 0-dB line, while the high-frequency asymptote [specified by eq. (6.27)] corresponds to a decrease of 20 dB in $|H(j\omega)|$ for every decade (i.e., factor of 10) in ω. This is sometimes referred to as a "20-dB-per-decade" asymptote.

Note that the two asymptotic approximations given in eqs. (6.26) and (6.27) are equal at the point $\log_{10}(\omega) = -\log_{10}(\tau)$, or equivalently, $\omega = 1/\tau$. Interpreted graphically, this means that the two straight-line asymptotes meet at $\omega = 1/\tau$, which suggests a straight-line approximation to the magnitude plot. That is, our approximation to $20\log_{10}|H(j\omega)|$ equals 0 for $\omega \le 1/\tau$ and is given by eq. (6.27) for $\omega \ge 1/\tau$. This approximation is also sketched (as a dashed line) in Figure 6.20. The point at which the slope of the approximation changes is precisely $\omega = 1/\tau$, which, for this reason, is often referred to as the *break frequency.* Also, note that at $\omega = 1/\tau$ the two terms [$(\omega\tau)^2$ and 1] in the argument of the logarithm in eq. (6.25) are equal. Thus, at this point, the actual value of the magnitude is

$$20\log_{10}\left|H\left(j\frac{1}{\tau}\right)\right| = -10\log_{10}(2) \simeq -3\text{ dB}. \tag{6.28}$$

Because of this, the point $\omega = 1/\tau$ is sometimes called the 3-dB point. From the figure, we see that only near the break frequency is there any significant error in the straight-line approximate Bode plot. Thus, if we wish to obtain a more accurate sketch of the Bode plot, we need only modify the approximation near the break frequency.

It is also possible to obtain a useful straight-line approximation to $\sphericalangle H(j\omega)$:

$$\begin{aligned} \sphericalangle H(j\omega) &= -\tan^{-1}(\omega\tau) \\ &\simeq \begin{cases} 0, & \omega \le 0.1/\tau \\ -(\pi/4)[\log_{10}(\omega\tau)+1], & 0.1/\tau \le \omega \le 10/\tau. \\ -\pi/2, & \omega \ge 10/\tau \end{cases} \end{aligned} \tag{6.29}$$

Note that this approximation decreases linearly (from 0 to $-\pi/2$) as a function of $\log_{10}(\omega)$ in the range

$$\frac{0.1}{\tau} \le \omega \le \frac{10}{\tau},$$

i.e., in the range from one decade below the break frequency to one decade above the break frequency. Also, zero is the correct asymptotic value of $\sphericalangle H(j\omega)$ for $\omega \ll 1/\tau$, and $-\pi/2$ is the correct asymptotic value of $\sphericalangle H(j\omega)$ for $\omega \gg 1/\tau$. Furthermore, the approximation agrees with the actual value of $\sphericalangle H(j\omega)$ at the break frequency $\omega = 1/\tau$, at which point

$$\sphericalangle H\left(j\frac{1}{\tau}\right) = -\frac{\pi}{4}. \tag{6.30}$$

This asymptotic approximation is also plotted in Figure 6.20, and from it we can see how, if desired, we can modify the straight-line approximation to obtain a more accurate sketch of $\sphericalangle H(j\omega)$.

From this first-order system, we can again see the inverse relationship between time and frequency. As we make τ smaller, we speed up the time response of the system [i.e., $h(t)$ becomes more compressed toward the origin, and the rise time of the step response is reduced] and we simultaneously make the break frequency large [i.e., $H(j\omega)$ becomes broader, since $|H(j\omega)| \simeq 1$ for a larger range of frequencies]. This can also be seen by multiplying the impulse response by τ and observing the relationship between $\tau h(t)$ and $H(j\omega)$:

$$\tau h(t) = e^{-t/\tau}u(t), \quad H(j\omega) = \frac{1}{j\omega\tau + 1}.$$

Thus, $\tau h(t)$ is a function of t/τ and $H(j\omega)$ is a function of $\omega\tau$, and from this we see that changing τ is essentially equivalent to a scaling in time and frequency.

6.5.2 Second-Order Continuous-Time Systems

The linear constant-coefficient differential equation for a second-order system is

$$\frac{d^2y(t)}{dt^2} + 2\zeta\omega_n\frac{dy(t)}{dt} + \omega_n^2 y(t) = \omega_n^2 x(t). \tag{6.31}$$

Equations of this type arise in many physical systems, including *RLC* circuits and mechanical systems, such as the one illustrated in Figure 6.21, composed of a spring, a mass, and a viscous damper or dashpot. In the figure, the input is the applied force $x(t)$ and the output is the displacement of the mass $y(t)$ from some equilibrium position at which the spring exerts no restoring force. The equation of motion for this system is

$$m\frac{d^2y(t)}{dt^2} = x(t) - ky(t) - b\frac{dy(t)}{dt},$$

or

$$\frac{d^2y(t)}{dt^2} + \left(\frac{b}{m}\right)\frac{dy(t)}{dt} + \left(\frac{k}{m}\right)y(t) = \frac{1}{m}x(t).$$

Comparing this to eq. (6.31), we see that if we identify

$$\omega_n = \sqrt{\frac{k}{m}} \tag{6.32}$$

and

$$\zeta = \frac{b}{2\sqrt{km}},$$

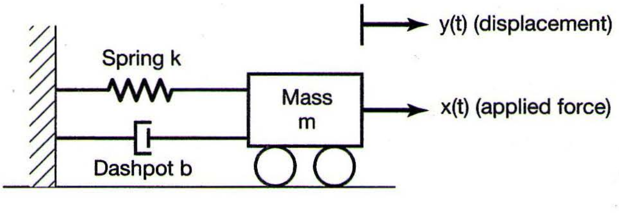


Figure 6.21 Second-order system consisting of a spring and dashpot attached to a moveable mass and a fixed support.

then [except for a scale factor of k on $x(t)$] the equation of motion for the system of Figure 6.21 reduces to eq. (6.31).

The frequency response for the second-order system of eq. (6.31) is

$$H(j\omega) = \frac{\omega_n^2}{(j\omega)^2 + 2\zeta\omega_n(j\omega) + \omega_n^2}. \tag{6.33}$$

The denominator of $H(j\omega)$ can be factored to yield

$$H(j\omega) = \frac{\omega_n^2}{(j\omega - c_1)(j\omega - c_2)},$$

where

$$\begin{aligned} c_1 &= -\zeta\omega_n + \omega_n\sqrt{\zeta^2 - 1}, \\ c_2 &= -\zeta\omega_n - \omega_n\sqrt{\zeta^2 - 1}. \end{aligned} \tag{6.34}$$

For $\zeta \neq 1$, c_1 and c_2 are unequal, and we can perform a partial-fraction expansion of the form

$$H(j\omega) = \frac{M}{j\omega - c_1} - \frac{M}{j\omega - c_2}, \tag{6.35}$$

where

$$M = \frac{\omega_n}{2\sqrt{\zeta^2 - 1}}. \tag{6.36}$$

From eq. (6.35), the corresponding impulse response for the system is

$$h(t) = M[e^{c_1 t} - e^{c_2 t}]u(t). \tag{6.37}$$

If $\zeta = 1$, then $c_1 = c_2 = -\omega_n$, and

$$H(j\omega) = \frac{\omega_n^2}{(j\omega + \omega_n)^2}. \tag{6.38}$$

From Table 4.2, we find that in this case the impulse response is

$$h(t) = \omega_n^2 t e^{-\omega_n t} u(t). \tag{6.39}$$

Note from eqs. (6.37) and (6.39), that $h(t)/\omega_n$ is a function of $\omega_n t$. Furthermore, eq. (6.33) can be rewritten as

$$H(j\omega) = \frac{1}{\left(j\omega/\omega_n\right)^2 + 2\zeta\left(j\omega/\omega_n\right) + 1},$$

from which we see that the frequency response is a function of ω/ω_n. Thus, changing ω_n is essentially identical to a time and frequency scaling.

The parameter ζ is referred to as the *damping ratio* and the parameter ω_n as the *undamped natural frequency.* The motivation for this terminology becomes clear when we take a more detailed look at the impulse response and the step response of a second-order system. First, from eq. (6.35), we see that for $0 < \zeta < 1$, c_1 and c_2 are complex, and we can rewrite the impulse response in eq. (6.37) in the form

$$\begin{aligned} h(t) &= \frac{\omega_n e^{-\zeta\omega_n t}}{2j\sqrt{1-\zeta^2}}\{\exp[j(\omega_n\sqrt{1-\zeta^2})t] - \exp[-j(\omega_n\sqrt{1-\zeta^2})t]\}u(t) \\ &= \frac{\omega_n e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}}[\sin(\omega_n\sqrt{1-\zeta^2})t]u(t). \end{aligned} \tag{6.40}$$

Thus, for $0 < \zeta < 1$, the second-order system has an impulse response that has damped oscillatory behavior, and in this case the system is referred to as being *under-damped.* If $\zeta > 1$, both c_1 and c_2 are real and negative, and the impulse response is the difference between two decaying exponentials. In this case, the system is *overdamped.* The case of $\zeta = 1$, when $c_1 = c_2$, is called the *critically damped* case. The impulse responses (multiplied by $1/\omega_n$) for second-order systems with different values of ζ are plotted in Figure 6.22(a).

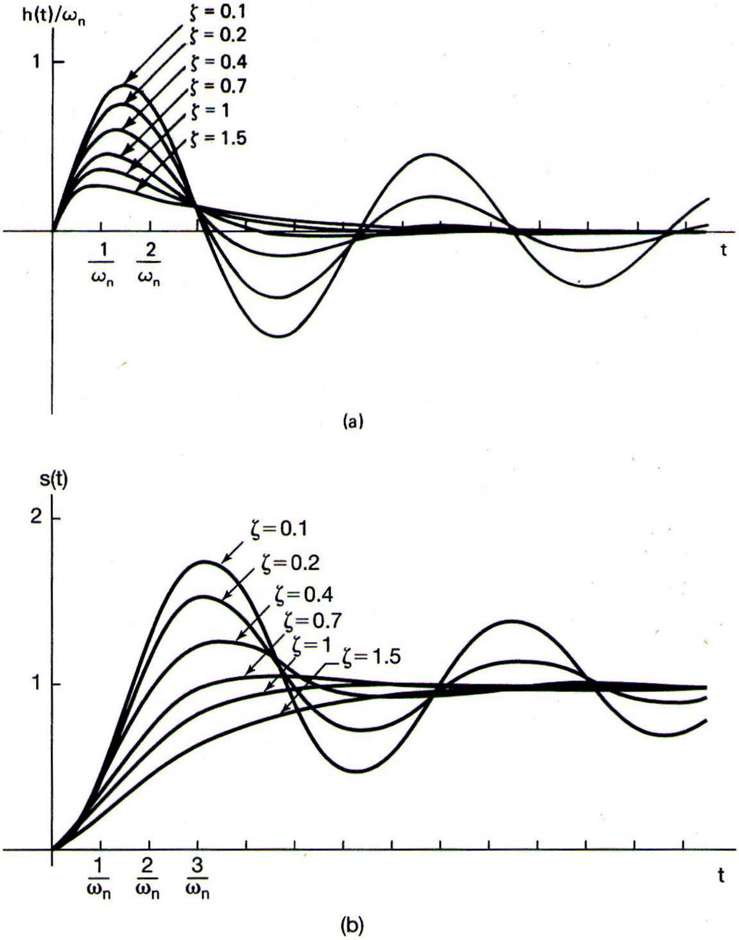


Figure 6.22 Response of continuous-time second-order systems with different values of the damping ratio ζ: (a) impulse response; (b) step response.

The step response of a second-order system can be calculated from eq. (6.37) for $\zeta \neq 1$. This yields the expression

$$s(t) = h(t) * u(t) = \left\{1 + M\left[\frac{e^{c_1 t}}{c_1} - \frac{e^{c_2 t}}{c_2}\right]\right\} u(t). \tag{6.41}$$

For $\zeta = 1$, we can use eq. (6.39) to obtain

$$s(t) = [1 - e^{-\omega_n t} - \omega_n t e^{-\omega_n t}]u(t). \tag{6.42}$$

The step response of a second-order system is plotted in Figure 6.22(b) for several values of ζ. From this figure, we see that in the underdamped case, the step response exhibits both *overshoot* (i.e., the step response exceeds its final value) and *ringing* (i.e., oscillatory behavior). For $\zeta = 1$, the step response has the fastest response (i.e., the shortest rise time) that is possible without overshoot and thus has the shortest settling time. As ζ increases beyond 1, the response becomes slower. This can be seen from eqs. (6.34) and (6.41). As ζ increases, c_1 becomes smaller in magnitude, while c_2 increases in magnitude. Therefore, although the time constant $(1/|c_2|)$ associated with $e^{c_2 t}$ decreases, the time constant $(1/|c_1|)$ associated with $e^{c_1 t}$ increases. Consequently the term involving $e^{c_1 t}$ in eq. (6.41) takes a longer time to decay to zero, and thus it is the time constant associated with this term that determines the settling time of the step response. As a result the step response takes longer to settle for large values of ζ. In terms of our spring-dashpot example, as we increase the magnitude of the damping coefficient b beyond the critical value at which ζ in eq. (6.33) equals 1, the motion of the mass becomes increasingly sluggish.

Finally, note that, as we have said, the value of ω_n essentially controls the time scale of the responses $h(t)$ and $s(t)$. For example, in the underdamped case, the larger ω_n is, the more compressed is the impulse response as a function of t, and the higher is the frequency of the oscillations or ringing in both $h(t)$ and $s(t)$. In fact, from eq. (6.40), we see that the frequency of the oscillations in $h(t)$ and $s(t)$ is $\omega_n\sqrt{1-\zeta^2}$, which does increase with increasing ω_n. Note, however, that this frequency depends explicitly on the damping ratio and does not equal (and is in fact smaller than) ω_n, except in the *undamped* case, $\zeta = 0$. (It is for this reason that the parameter ω_n is traditionally referred to as the undamped natural frequency.) For the spring-dashpot example, we therefore conclude that the rate of oscillation of the mass equals ω_n when no dashpot is present, and the oscillation frequency decreases when we include the dashpot.

In Figure 6.23, we have depicted the Bode plot of the frequency response given in eq. (6.33) for several values of ζ. As in the first-order case, the logarithmic frequency scale leads to linear high- and low-frequency asymptotes for the log magnitude. Specifically, from eq. (6.33),

$$20\log_{10}|H(j\omega)| = -10\log_{10}\left\{\left[1 - \left(\frac{\omega}{\omega_n}\right)^2\right]^2 + 4\zeta^2\left(\frac{\omega}{\omega_n}\right)^2\right\}. \tag{6.43}$$

From this expression, it follows that

$$20\log_{10}|H(j\omega)| \simeq \begin{cases} 0, & \text{for } \omega \ll \omega_n \\ -40\log_{10}\omega + 40\log_{10}\omega_n, & \text{for } \omega \gg \omega_n \end{cases}. \tag{6.44}$$

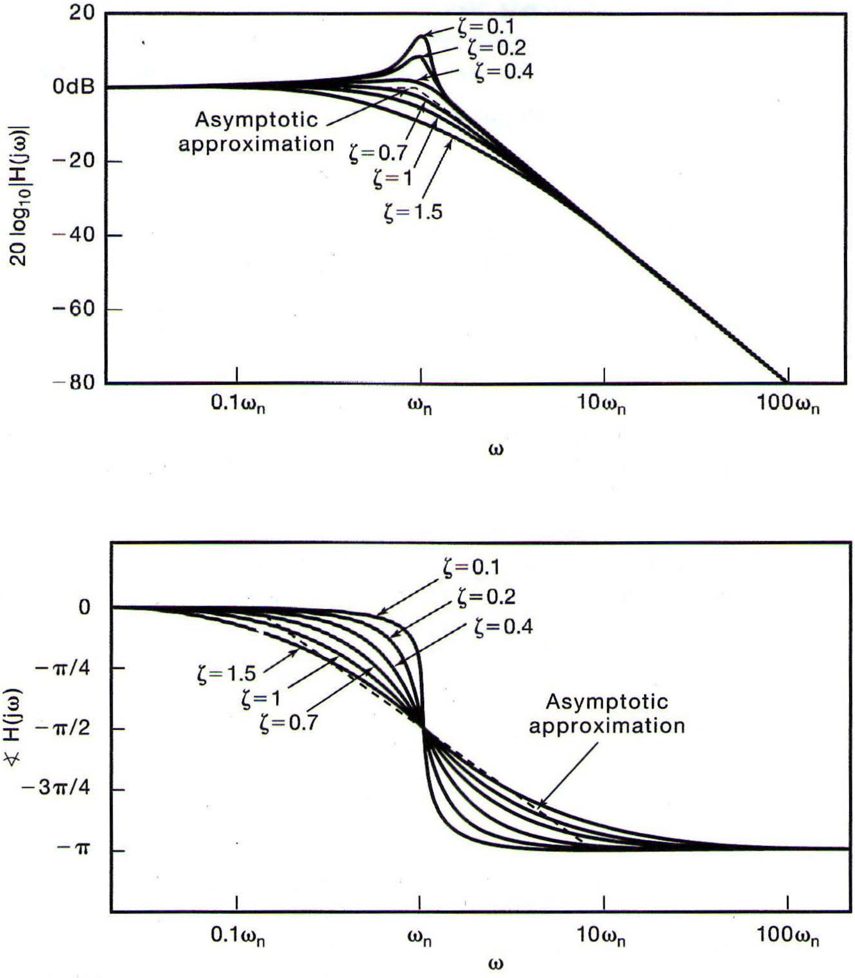


Figure 6.23 Bode plots for second-order systems with several different values of damping ratio ζ.

Therefore, the low-frequency asymptote of the log magnitude is the 0-dB line, while the high-frequency asymptote [given by eq. (6.44)] has a slope of -40 dB per decade; i.e., $|H(j\omega)|$ decreases by 40 dB for every increase in ω of a factor of 10. Also, note that the two straight-line asymptotes meet at the point $\omega = \omega_n$. Thus, we obtain a straight-line approximation to the log magnitude by using the approximation given in eq. (6.44) for $\omega \le \omega_n$. For this reason, ω_n is referred to as the break frequency of the second-order system. This approximation is also plotted (as a dashed line) in Figure 6.23.

We can, in addition, obtain a straight-line approximation to $\sphericalangle H(j\omega)$, whose exact expression can be obtained from eq. (6.33):

$$\sphericalangle H(j\omega) = -\tan^{-1}\left(\frac{2\zeta(\omega/\omega_n)}{1-(\omega/\omega_n)^2}\right). \tag{6.45}$$

The approximation is

$$\sphericalangle H(j\omega) \simeq \begin{cases} 0, & \omega \le 0.1\omega_n \\ -\frac{\pi}{2}\left[\log_{10}\left(\frac{\omega}{\omega_n}\right) + 1\right], & 0.1\omega_n \le \omega \le 10\omega_n, \\ -\pi, & \omega \ge 10\omega_n \end{cases} \tag{6.46}$$

which is also plotted in Figure 6.23. Note that the approximation and the actual value again are equal at the break frequency $\omega = \omega_n$, where

$$\sphericalangle H(j\omega_n) = -\frac{\pi}{2}.$$

It is important to observe that the asymptotic approximations, eqs. (6.44) and (6.46), we have obtained for a second-order system do not depend on ζ, while the actual plots of $|H(j\omega)|$ and $\sphericalangle H(j\omega)$ certainly do, and thus, to obtain an accurate sketch, especially near the break frequency $\omega = \omega_n$, we must take this into account by modifying the approximations to conform more closely to the actual plots. The discrepancy is most pronounced for small values of ζ. In particular, note that in this case the actual log magnitude has a peak around $\omega = \omega_n$. In fact, straightforward calculations using eq. (6.43) show that, for $\zeta < \sqrt{2}/2 \simeq 0.707$, $|H(j\omega)|$ has a maximum value at

$$\omega_{\max} = \omega_n\sqrt{1 - 2\zeta^2}, \tag{6.47}$$

and the value at this maximum point is

$$|H(j\omega_{\max})| = \frac{1}{2\zeta\sqrt{1-\zeta^2}}. \tag{6.48}$$

For $\zeta > 0.707$, however, $H(j\omega)$ decreases monotonically as ω increases from zero. The fact that $H(j\omega)$ can have a peak is extremely important in the design of frequency-selective filters and amplifiers. In some applications, one may want to design such a circuit so that it has a sharp peak in the magnitude of its frequency response at some specified frequency, thereby providing large frequency-selective amplification for sinusoids at frequencies within a narrow band. The *quality* Q of such a circuit is defined to be a measure of the sharpness of the peak. For a second-order circuit described by an equation of the form of eq. (6.31), the quality is usually taken to be

$$Q = \frac{1}{2\zeta},$$

and from Figure 6.23 and eq. (6.48), we see that this definition has the proper behavior: The less damping there is in the system, the sharper is the peak in $|H(j\omega)|$.

6.5.3 Bode Plots for Rational Frequency Responses

At the start of this section, we indicated that first- and second-order systems can be used as basic building blocks for more complex LTI systems with rational frequency responses. One consequence of this is that the Bode plots presented here essentially provide us with all of the information we need to construct Bode plots for arbitrary rational frequency responses. Specifically, we have described the Bode plots for the frequency responses given by eqs. (6.22) and (6.33). In addition, we can readily obtain the Bode plots for frequency responses of the forms

$$H(j\omega) = 1 + j\omega\tau \tag{6.49}$$

and

$$H(j\omega) = 1 + 2\zeta\left(\frac{j\omega}{\omega_n}\right) + \left(\frac{j\omega}{\omega_n}\right)^2. \tag{6.50}$$

The Bode plots for eqs. (6.49) and (6.50) follow directly from Figures 6.20 and 6.23 and from the fact that

$$20\log_{10}|H(j\omega)| = -20\log_{10}\left|\frac{1}{H(j\omega)}\right|$$

and

$$\sphericalangle(H(j\omega)) = -\sphericalangle\left(\frac{1}{H(j\omega)}\right).$$

Also, consider a system function that is a constant gain

$$H(j\omega) = K.$$

Since $K = |K|e^{j\cdot 0}$ if $K > 0$ and $K = |K|e^{j\pi}$ if $K < 0$, we see that

$$20\log_{10}|H(j\omega)| = 20\log_{10}|K|$$

$$\sphericalangle H(j\omega) = \begin{cases} 0, & \text{if } K > 0 \\ \pi, & \text{if } K < 0 \end{cases}$$

Since a rational frequency response can be factored into the product of a constant gain and first- and second-order terms, its Bode plot can be obtained by summing the plots for each of the terms. We illustrate further the construction of Bode plots in the next two examples.

Example 6.4

Let us obtain the Bode plot for the frequency response

$$H(j\omega) = \frac{2\times 10^4}{(j\omega)^2 + 100j\omega + 10^4}.$$

First, we note that

$$H(j\omega) = 2\hat{H}(j\omega),$$

where $\hat{H}(j\omega)$ has the same form as the standard second-order frequency response specified by eq. (6.33). It follows that

$$20\log_{10}|H(j\omega)| = 20\log_{10}2 + 20\log|\hat{H}(j\omega)|.$$

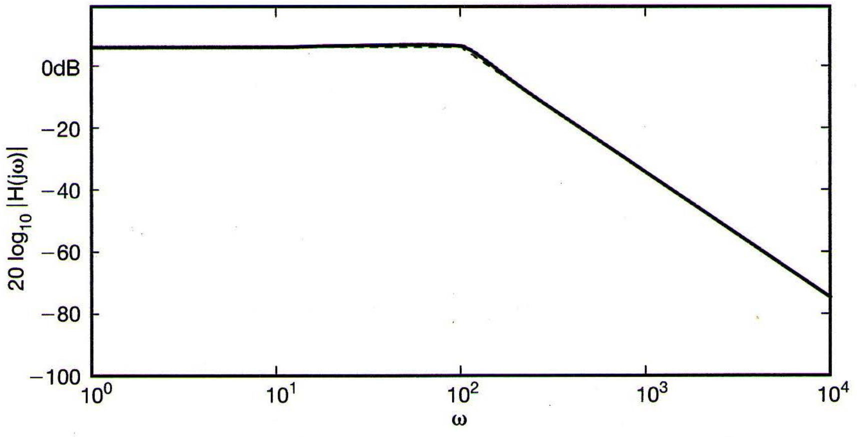
By comparing $\hat{H}(j\omega)$ with the frequency response in eq. (6.33), we conclude that $\omega_n = 100$ and $\zeta = 1/2$ for $\hat{H}(j\omega)$. Using eq. (6.44), we may now specify the asymptotes for $20\log_{10}|\hat{H}(j\omega)|$:

$$20\log_{10}|\hat{H}(j\omega)| \simeq 0 \quad \text{for } \omega \ll 100,$$

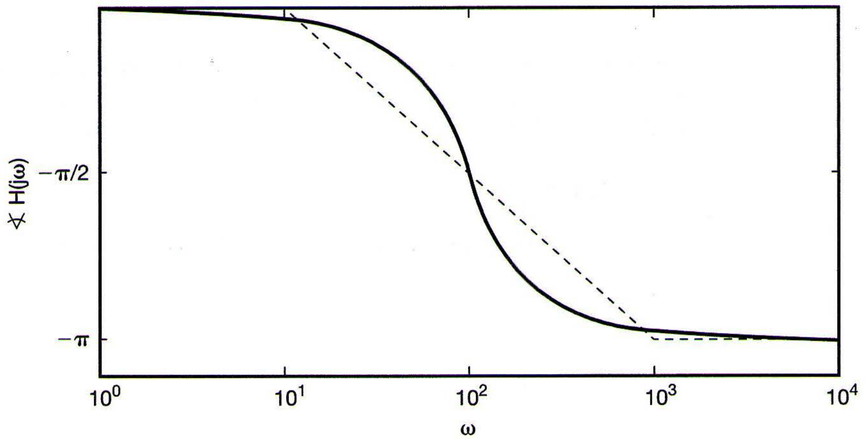
and

$$20\log_{10}|\hat{H}(j\omega)| \simeq -40\log_{10}\omega + 80 \quad \text{for } \omega \gg 100.$$

It follows that $20\log_{10}|H(j\omega)|$ will have the same asymptotes, except for a constant offset at all frequencies due to the addition of the $20\log_{10}2$ term (which approximately equals 6 dB). The dashed lines in Figure 6.24(a) represent these asymptotes.



(a)



(b)

Figure 6.24 Bode plot for system function in Example 6.4: (a) magnitude; (b) phase.

The solid curve in the same figure represents the actual computer-generated Bode plot for $20\log_{10}|H(j\omega)|$. Since the value of ζ for $\hat{H}(j\omega)$ is less than $\sqrt{2}/2$, the actual Bode plot has a slight peak near $\omega = 100$.

To obtain a plot of $\sphericalangle H(j\omega)$, we note that

$$\sphericalangle H(j\omega) = \sphericalangle \hat{H}(j\omega)$$

and that $\sphericalangle \hat{H}(j\omega)$ has its asymptotes specified in accordance with eq. (6.46); that is,

$$\sphericalangle \hat{H}(j\omega) = \begin{cases} 0, & \omega \leq 10 \\ -(\pi/2)[\log_{10}(\omega/100) + 1], & 10 \leq \omega \leq 1{,}000. \\ -\pi, & \omega \geq 1{,}000. \end{cases}$$

The asymptotes and the actual values for $\sphericalangle H(j\omega)$ are plotted with dashed and solid lines, respectively, in Figure 6.24(b).

Example 6.5

Consider the frequency response

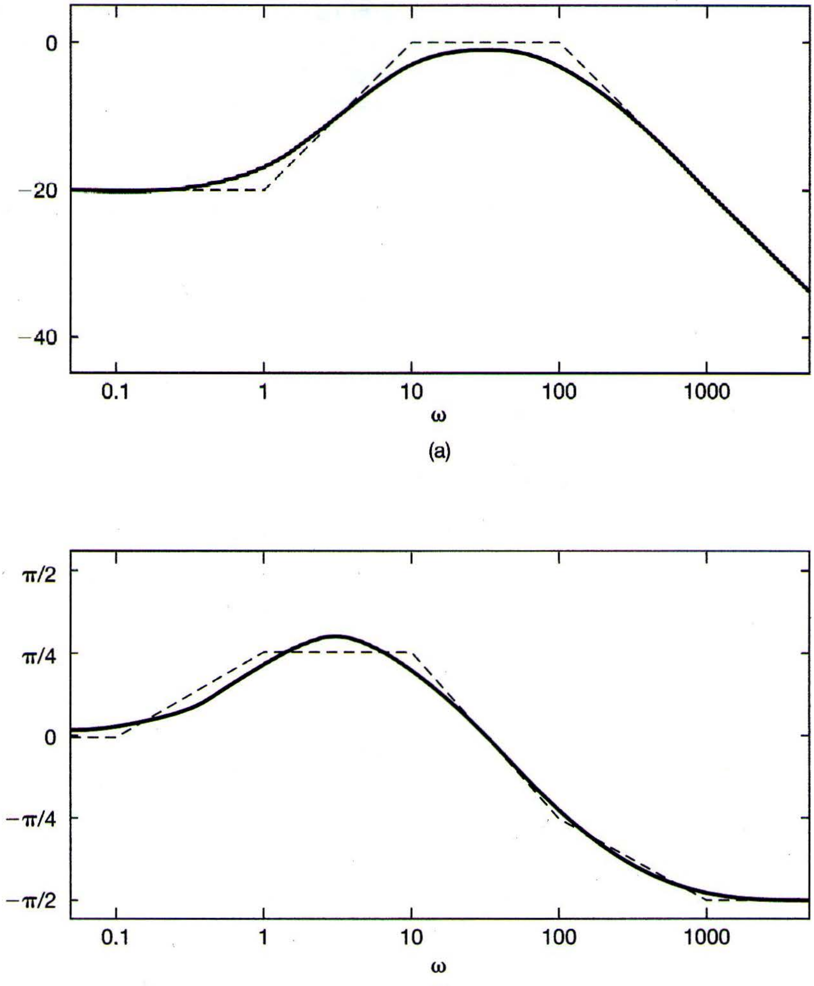
$$H(j\omega) = \frac{100(1 + j\omega)}{(10 + j\omega)(100 + j\omega)}.$$

To obtain the Bode plot for $H(j\omega)$, we rewrite it in the following factored form:

$$H(j\omega) = \left(\frac{1}{10}\right)\left(\frac{1}{1 + j\omega/10}\right)\left(\frac{1}{1 + j\omega/100}\right)(1 + j\omega).$$

Here, the first factor is a constant, the next two factors have the standard form for a first-order frequency response as specified in eq. (6.22), and the fourth factor is the reciprocal of the same first-order standard form. The Bode plot for $20\log_{10}|H(j\omega)|$ is therefore the sum of the Bode plots corresponding to each of the factors. Furthermore, the asymptotes corresponding to each factor may be summed to obtain the asymptotes for the overall Bode plot. These asymptotes and the actual values of $20\log_{10}|H(j\omega)|$ are displayed in Figure 6.25(a). Note that the constant factor of 1/10 accounts for an offset of −20 dB at each frequency. The break frequency at $\omega = 1$ corresponds to the $(1 + j\omega)$ factor, which produces the 20 dB/decade rise that starts at $\omega = 1$ and is canceled by the 20 dB/decade decay that starts at the break frequency at $\omega = 10$ and is due to the $1/(1 + j\omega/10)$ factor. Finally, the $1/(1 + j\omega/100)$ factor contributes another break frequency at $\omega = 100$ and a subsequent decay at the rate of 20 dB/decade.

Similarly we can construct the asymptotic approximation for $\sphericalangle H(j\omega)$ from the individual asymptotes for each factor, as illustrated, together with a plot of the exact value of the phase, in Figure 6.25(b). In particular, the constant factor 1/10 contributes 0 to the phase, while the factor $(1 + j\omega)$ contributes an asymptotic approximation that is 0 for $\omega < 0.1$, and rises linearly as a function of $\log_{10}(\omega)$ from a value of zero at $\omega = 0.1$ to a value of $\pi/2$ radians at $\omega = 10$. However, this rise is canceled at $\omega = 1$ by the asymptotic approximation for the angle of $1/(1 + j\omega/10)$ which contributes a linear decrease in angle of $\pi/2$ radians over the range of frequencies from $\omega = 1$ to $\omega = 100$. Finally, the asymptotic approximation for the angle of $1/(1 + j\omega/100)$ contributes another linear decrease in angle of $\pi/2$ radians over the range of frequencies from $\omega = 10$ to $\omega = 1000$.



(a)

(b)

Figure 6.25 Bode plot for system function in Example 6.5: (a) magnitude; (b) phase.

In our discussion of first-order systems in this section, we restricted our attention to values of $\tau > 0$. In fact, it is not difficult to check that if $\tau < 0$, then the causal first-order system described by eq. (6.21) has an impulse response that is not absolutely integrable, and consequently, the system is unstable. Similarly, in analyzing the second-order causal system in eq. (6.31), we required that both ζ and ω_n^2 be positive numbers. If either of these is not positive, the resulting impulse response is not absolutely integrable. Thus, in this section we have restricted attention to those causal first- and second-order systems that are stable and for which we can define frequency responses.

6.6 FIRST-ORDER AND SECOND-ORDER DISCRETE-TIME SYSTEMS

In this section, we examine the properties of first- and second-order discrete-time LTI systems, paralleling the development in the preceding section. As in continuous time, any system with a frequency response that is a ratio of polynomials in $e^{-j\omega}$—i.e., any discrete-time LTI system described by a linear constant-coefficient difference equation—can be written as a product or sum of first- and second-order systems, implying that these basic systems are of considerable value in both implementing and analyzing more complex systems. (See, for example, Problem 6.45.)

6.6.1 First-Order Discrete-Time Systems

Consider the first-order causal LTI system described by the difference equation

$$y[n] - ay[n-1] = x[n], \tag{6.51}$$

with $|a| < 1$. From Example 5.18, the frequency response of this system is

$$H(e^{j\omega}) = \frac{1}{1 - ae^{-j\omega}}, \tag{6.52}$$

and its impulse response is

$$h[n] = a^n u[n], \tag{6.53}$$

which is illustrated in Figure 6.26 for several values of a. Also, the step response of the system is

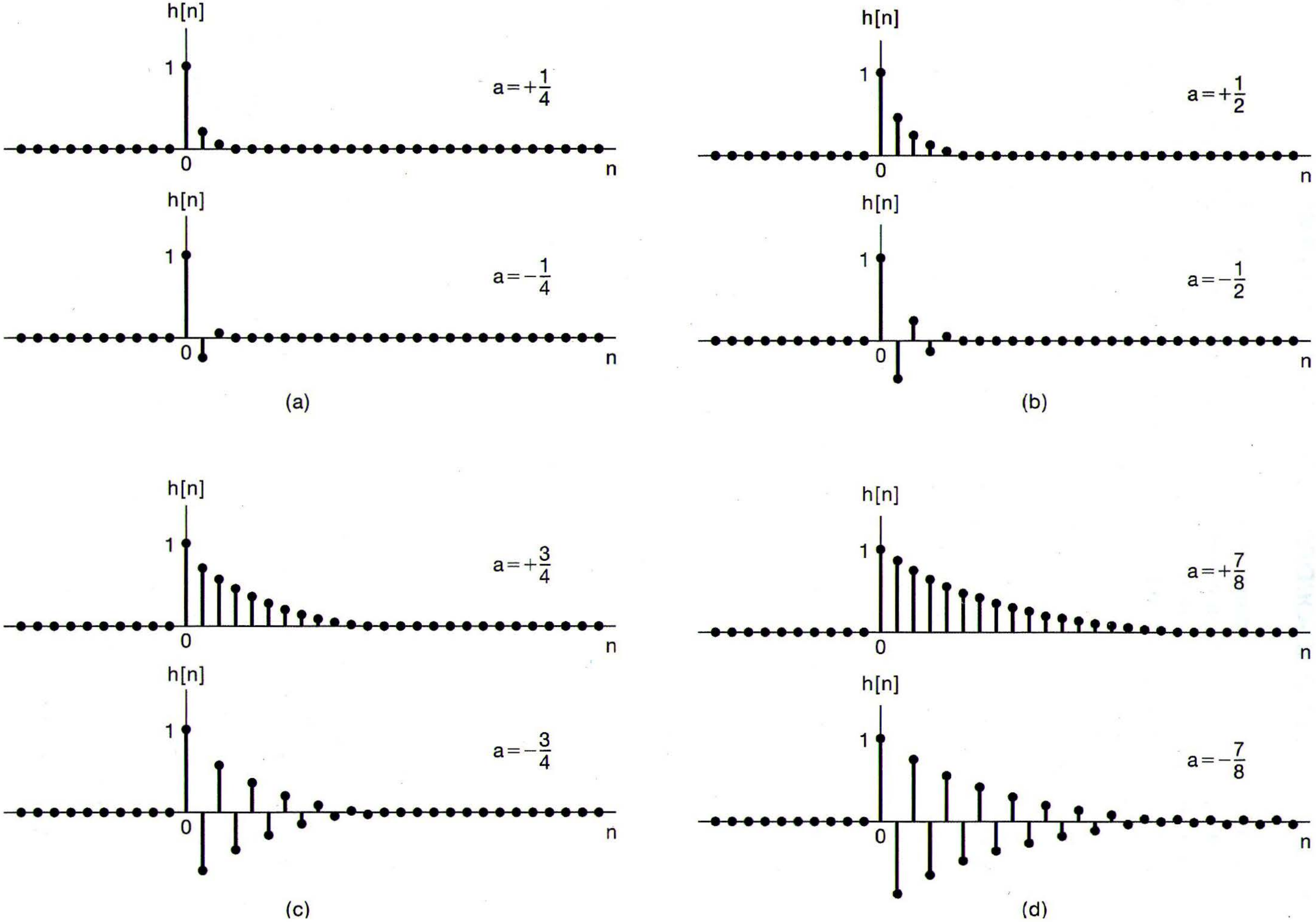
$$s[n] = h[n] * u[n] = \frac{1 - a^{n+1}}{1 - a} u[n], \tag{6.54}$$

which is illustrated in Figure 6.27.

The magnitude of the parameter a plays a role similar to that of the time constant τ in the continuous-time first-order system. Specifically, $|a|$ determines the rate at which the first-order system responds. For example, from eqs. (6.53) and (6.54) and Figures 6.26 and 6.27, we see that $h[n]$ and $s[n]$ converge to their final value at the rate at which $|a|^n$ converges to zero. Therefore, the impulse response decays sharply and the step response settles quickly for $|a|$ small. For $|a|$ nearer to 1, these responses are slower. Note that unlike its continuous-time counterpart, the first-order system described by eq. (6.51) can display oscillatory behavior. This occurs when $a < 0$, in which case the step response exhibits both overshoot of its final value and ringing.

The magnitude and phase of the frequency response of the first-order system in eq. (6.51) are, respectively,

$$|H(e^{j\omega})| = \frac{1}{(1 + a^2 - 2a\cos\omega)^{1/2}} \tag{6.55}$$



(a) (b) (c) (d)

Figure 6.26 Impulse response $h[n] = a^n u[n]$ of a first-order system: (a) $a = \pm 1/4$; (b) $a = \pm 1/2$; (c) $a = \pm 3/4$; (d) $a = \pm 7/8$.

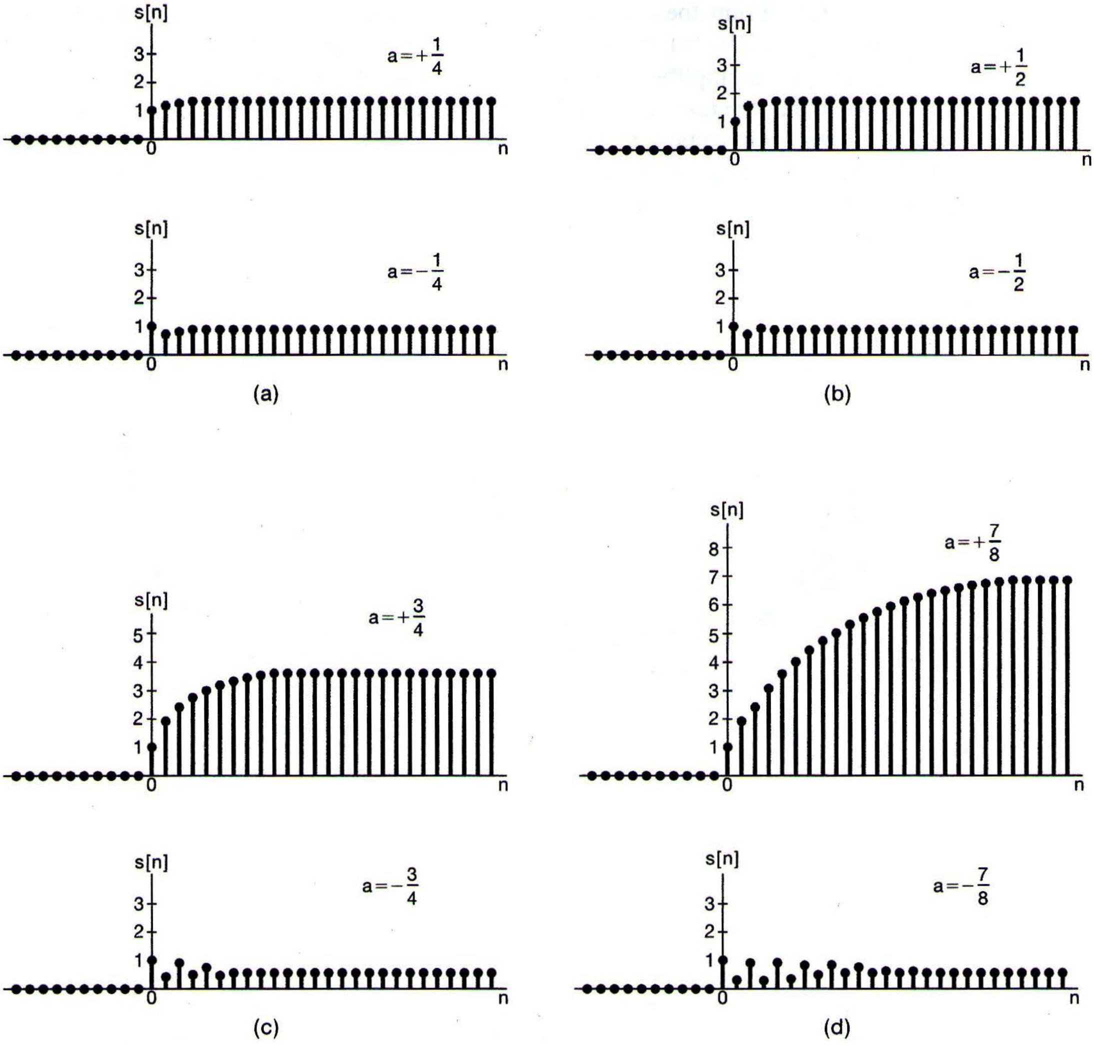


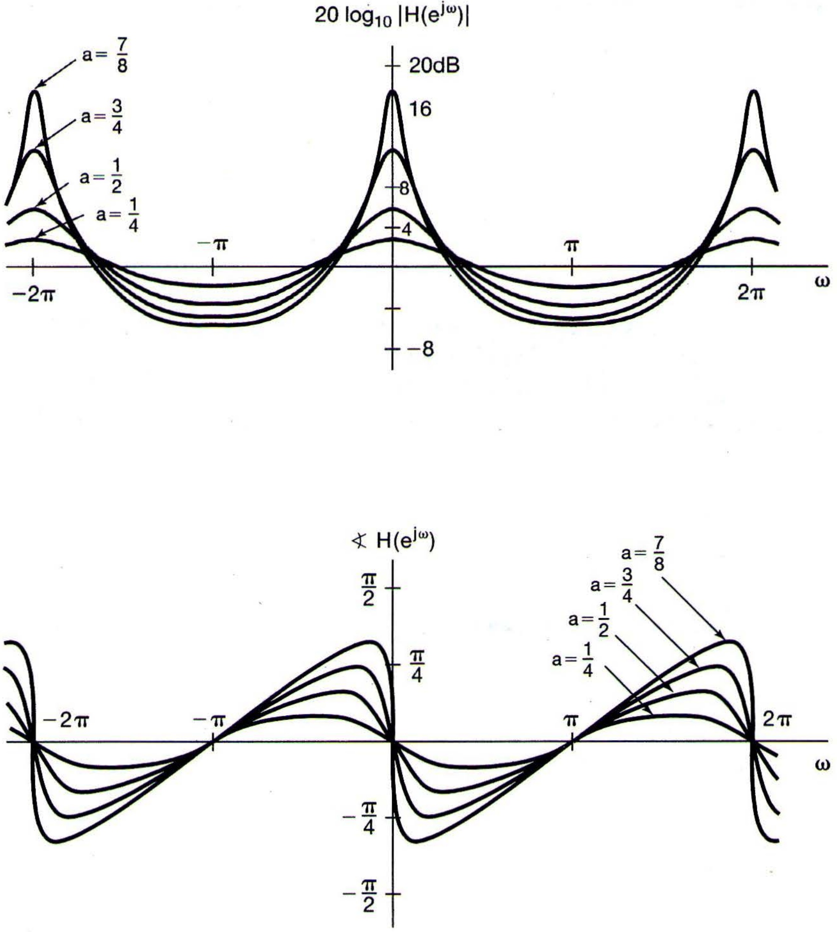
Figure 6.27 Step response $s[n]$ of a first-order system: (a) $a = \pm 1/4$; (b) $a = \pm 1/2$; (c) $a = \pm 3/4$; (d) $a = \pm 7/8$.

and

$$\sphericalangle H(e^{j\omega}) = -\tan^{-1}\left[\frac{a \sin\omega}{1 - a\cos\omega}\right]. \tag{6.56}$$

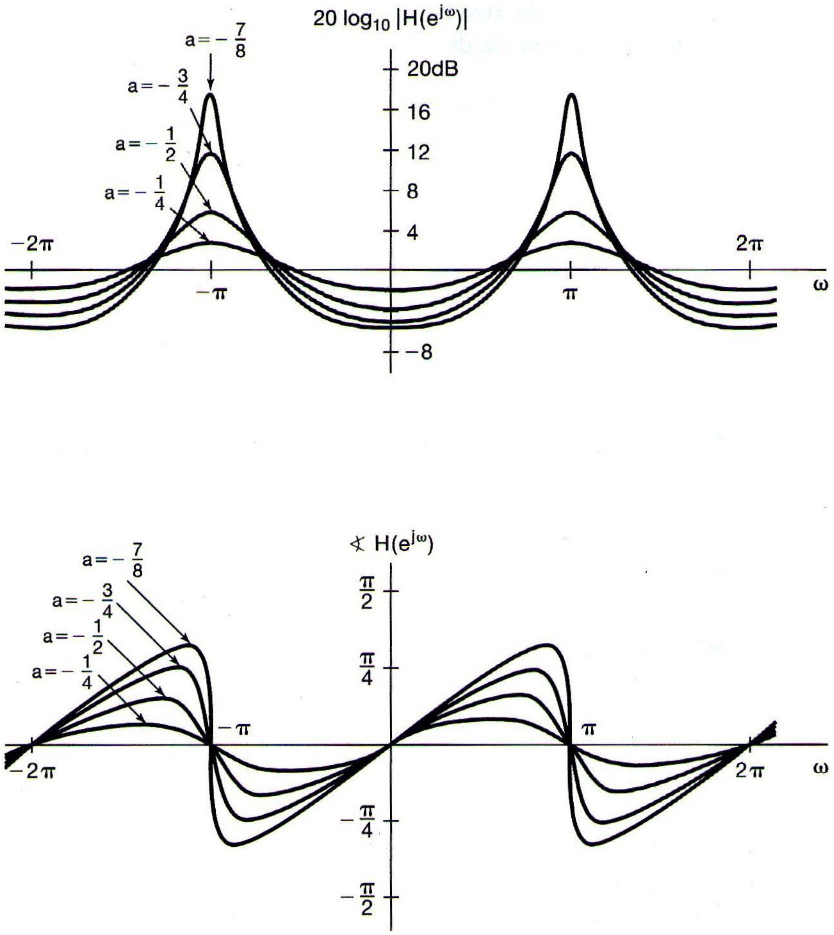
In Figure 6.28(a), we have plotted the log magnitude and the phase of the frequency response in eq. (6.52) for several values of $a > 0$. The case of $a < 0$ is illustrated in

Figure 6.28(b). From these figures, we see that for $a > 0$, the system attenuates high frequencies [i.e., $|H(e^{j\omega})|$ is smaller for ω near $\pm\pi$ than it is for ω near 0], while when $a < 0$, the system amplifies high frequencies and attenuates low frequencies. Note also that for $|a|$ small, the maximum and minimum values, $1/(1+a)$ and $1/(1-a)$, of $|H(e^{j\omega})|$ are close together in value, and the graph of $|H(e^{j\omega})|$ is relatively flat. On the other hand, for $|a|$ near 1, these quantities differ significantly, and consequently $|H(e^{j\omega})|$ is more sharply peaked, providing filtering and amplification that is more selective over a narrow band of frequencies.



(a)

Figure 6.28 Magnitude and phase of the frequency response of eq. (6.52) for a first-order system: (a) plots for several values of $a > 0$; (b) plots for several values of $a < 0$.



(b)

Figure 6.28 *Continued*

6.6.2 Second-Order Discrete-Time Systems

Consider next the second-order causal LTI system described by

$$y[n] - 2r\cos\theta\, y[n-1] + r^2 y[n-2] = x[n], \tag{6.57}$$

with $0 < r < 1$ and $0 \le \theta \le \pi$. The frequency response for this system is

$$H(e^{j\omega}) = \frac{1}{1 - 2r\cos\theta\, e^{-j\omega} + r^2 e^{-j2\omega}}. \tag{6.58}$$

The denominator of $H(e^{j\omega})$ can be factored to obtain

$$H(e^{j\omega}) = \frac{1}{[1 - (re^{j\theta})e^{-j\omega}][1 - (re^{-j\theta})e^{-j\omega}]}. \tag{6.59}$$

For $\theta \neq 0$ or π, the two factors in the denominator of $H(e^{j\omega})$ are different, and a partial-fraction expansion yields

$$H(e^{j\omega}) = \frac{A}{1-(re^{j\theta})e^{-j\omega}} + \frac{B}{1-(re^{-j\theta})e^{-j\omega}}, \tag{6.60}$$

where

$$A = \frac{e^{j\theta}}{2j\sin\theta}, \quad B = \frac{e^{-j\theta}}{2j\sin\theta}. \tag{6.61}$$

In this case, the impulse response of the system is

$$\begin{aligned} h[n] &= [A(re^{j\theta})^n + B(re^{-j\theta})^n]u[n] \\ &= r^n\frac{\sin[(n+1)\theta]}{\sin\theta}u[n]. \end{aligned} \tag{6.62}$$

For $\theta = 0$ or π, the two factors in the denominator of eq. (6.58) are the same. When $\theta = 0$,

$$H(e^{j\omega}) = \frac{1}{(1-re^{-j\omega})^2} \tag{6.63}$$

and

$$h[n] = (n+1)r^n u[n]. \tag{6.64}$$

When $\theta = \pi$,

$$H(e^{j\omega}) = \frac{1}{(1+re^{-j\omega})^2} \tag{6.65}$$

and

$$h[n] = (n+1)(-r)^n u[n]. \tag{6.66}$$

The impulse responses for second-order systems are plotted in Figure 6.29 for a range of values of r and θ. From this figure and from eq. (6.62), we see that the rate of decay of $h[n]$ is controlled by r—i.e., the closer r is to 1, the slower is the decay in $h[n]$. Similarly, the value of θ determines the frequency of oscillation. For example, with $\theta = 0$ there is no oscillation in $h[n]$, while for $\theta = \pi$ the oscillations are rapid. The effect of different values of r and θ can also be seen by examining the step response of eq. (6.57). For $\theta \neq 0$ or π,

$$s[n] = h[n] * u[n] = \left[A\left(\frac{1-(re^{j\theta})^{n+1}}{1-re^{j\theta}}\right) + B\left(\frac{1-(re^{-j\theta})^{n+1}}{1-re^{-j\theta}}\right)\right]u[n]. \tag{6.67}$$

Also, using the result of Problem 2.52, we find that for $\theta = 0$,

$$s[n] = \left[\frac{1}{(r-1)^2} - \frac{r}{(r-1)^2}r^n + \frac{r}{r-1}(n+1)r^n\right]u[n], \tag{6.68}$$

while for $\theta = \pi$,

$$s[n] = \left[\frac{1}{(r+1)^2} + \frac{r}{(r+1)^2}(-r)^n + \frac{r}{r+1}(n+1)(-r)^n\right]u[n]. \tag{6.69}$$

The step response is plotted in Figure 6.30, again for a range of values of r and θ.

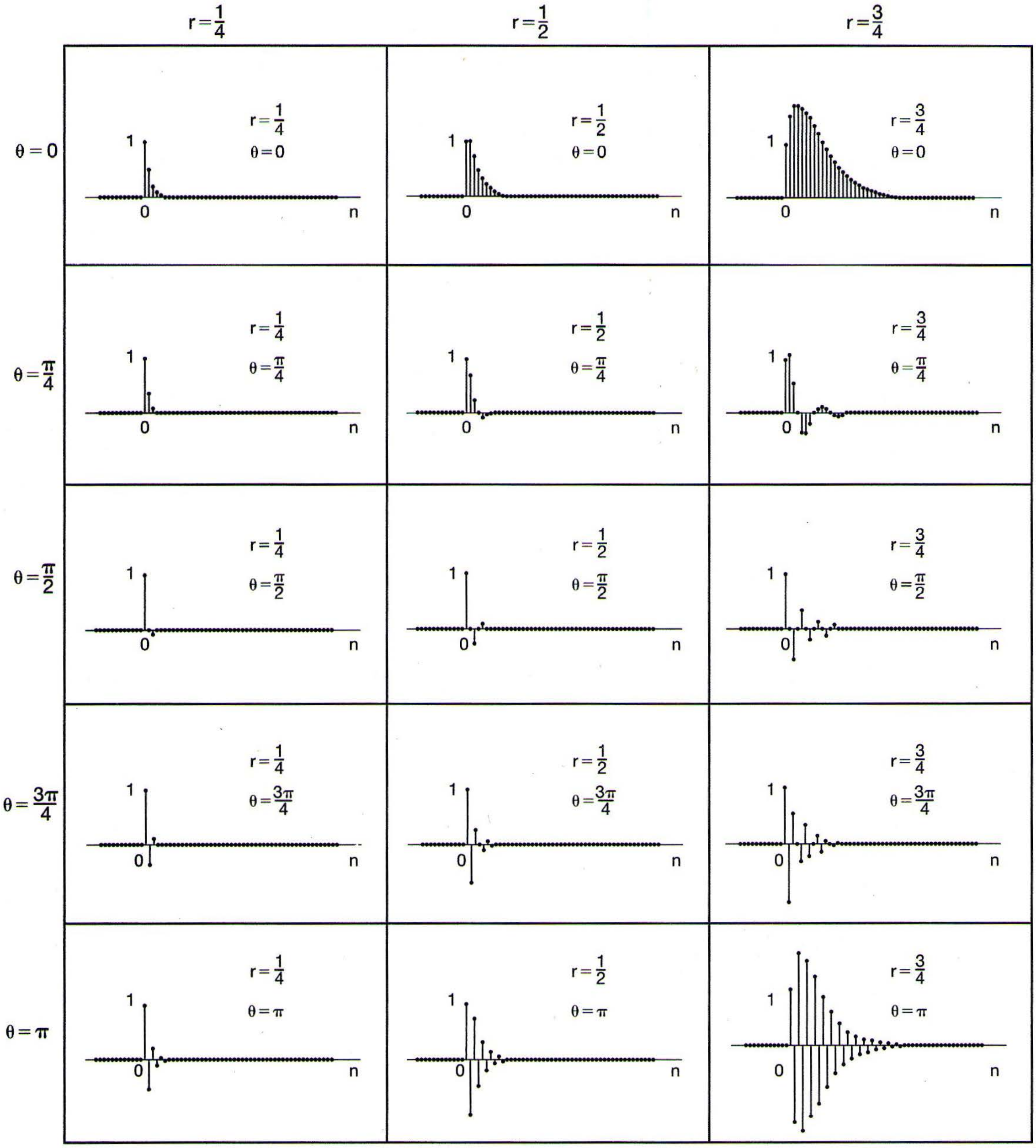
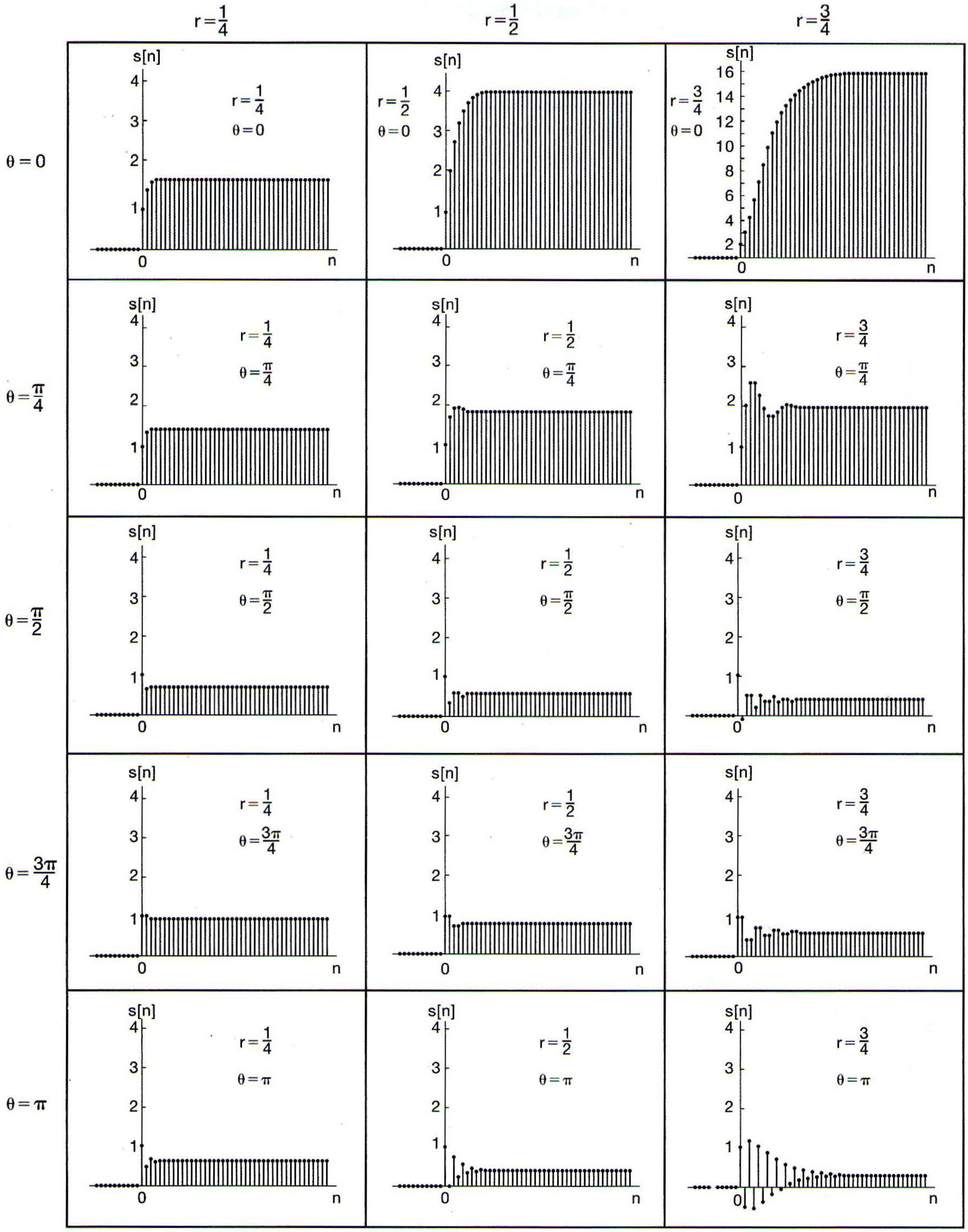


Figure 6.29 Impulse response of the second-order system of eq. (6.57) for a range of values of r and θ.

The second-order system given by eq. (6.57) is the counterpart of the *underdamped* second-order system in continuous time, while the special case of $\theta = 0$ is the critically damped case. That is, for any value of θ other than zero, the impulse response has a damped



* Note: The plot for $r=\frac{3}{4}$, $\theta=0$ has a different scale from the others.

Figure 6.30 Step response of the second-order system of eq. (6.57) for a range of values of r and θ.

oscillatory behavior, and the step response exhibits ringing and overshoot. The frequency response of this system is depicted in Figure 6.31 for a number of values of r and θ. From Figure 6.31, we see that a band of frequencies is amplified , and r determines how sharply peaked the frequency response is within this band.

As we have just seen, the second-order system described in eq. (6.59) has factors with complex coefficients (unless $\theta = 0$ or π). It is also possible to consider second-order systems having factors with real coefficients. Specifically, consider

$$H(e^{j\omega}) = \frac{1}{(1 - d_1 e^{-j\omega})(1 - d_2 e^{-j\omega})}, \tag{6.70}$$

where d_1 and d_2 are both real numbers with $|d_1|, |d_2| < 1$. Equation (6.70) is the frequency response for the difference equation

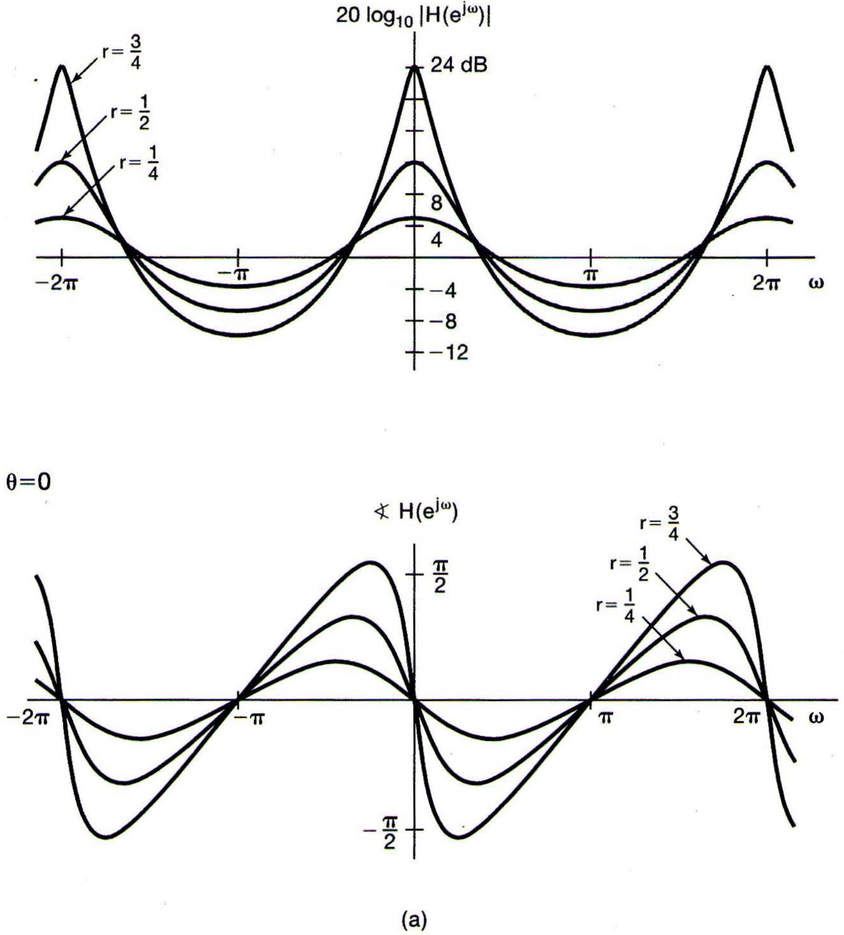
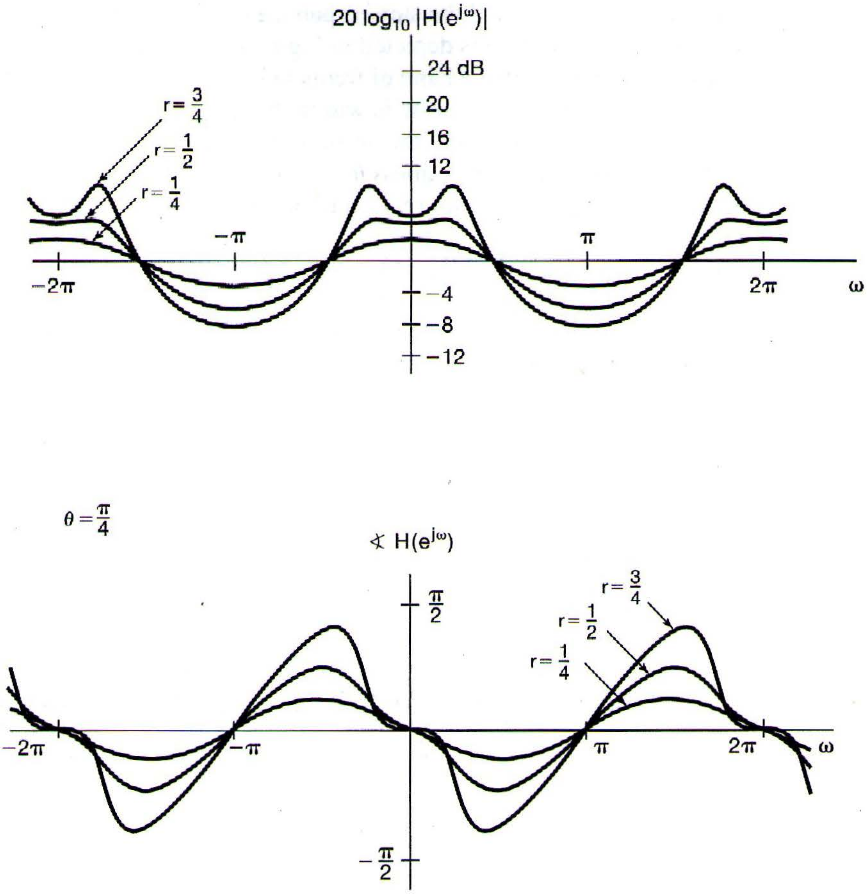


Figure 6.31 Magnitude and phase of the frequency response of the second-order system of eq. (6.57): (a) $\theta = 0$; (b) $\theta = \pi/4$; (c) $\theta = \pi/2$; (d) $\theta = 3\pi/4$; (e) $\theta = \pi$. Each plot contains curves corresponding to $r = 1/4$, $1/2$, and $3/4$.



(b)

Figure 6.31 *Continued*

$$y[n] - (d_1 + d_2)y[n-1] + d_1 d_2 y[n-2] = x[n]. \tag{6.71}$$

In this case,

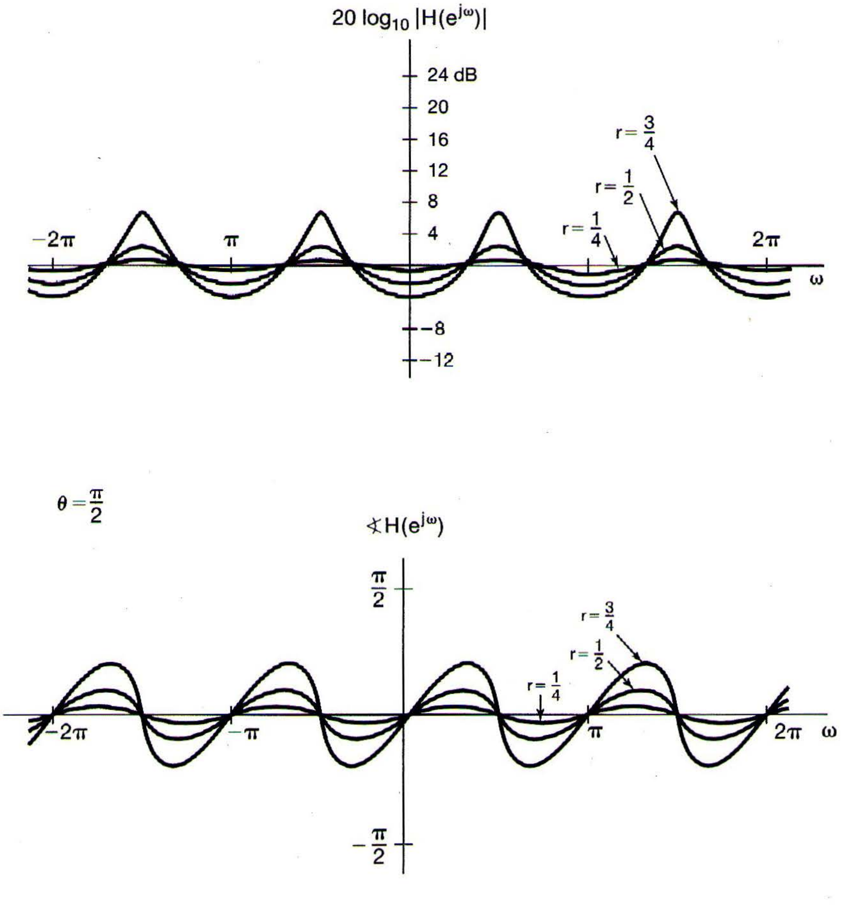
$$H(e^{j\omega}) = \frac{A}{1 - d_1 e^{-j\omega}} + \frac{B}{1 - d_2 e^{-j\omega}}, \tag{6.72}$$

where

$$A = \frac{d_1}{d_1 - d_2}, \quad B = \frac{d_2}{d_2 - d_1}. \tag{6.73}$$

Thus,

$$h[n] = [A d_1^n + B d_2^n]u[n], \tag{6.74}$$



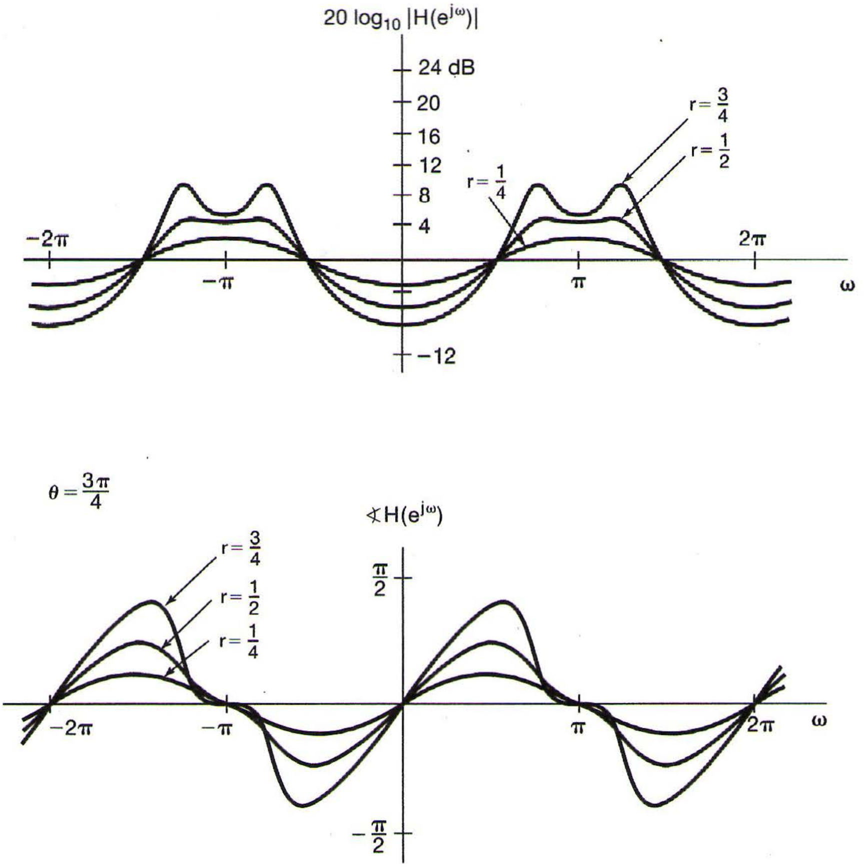
(c)

Figure 6.31 *Continued*

which is the sum of two decaying real exponentials. Also,

$$s[n] = \left[A\left(\frac{1-d_1^{n+1}}{1-d_1}\right) + B\left(\frac{1-d_2^{n+1}}{1-d_2}\right)\right]u[n]. \tag{6.75}$$

The system with frequency response given by eq. (6.70) corresponds to the cascade of two first-order systems. Therefore, we can deduce most of its properties from our understanding of the first-order case. For example, the log-magnitude and phase plots for eq. (6.70) can be obtained by adding together the plots for each of the two first-order terms. Also, as we saw for first-order systems, the response of the system is fast if $|d_1|$ and $|d_2|$ are small, but the system has a long settling time if either of these magnitudes is near 1. Furthermore, if d_1 and d_2 are negative, the response is oscillatory. The case when both d_1 and d_2 are positive is the counterpart of the overdamped case in continuous time, with the impulse and step responses settling without oscillation.



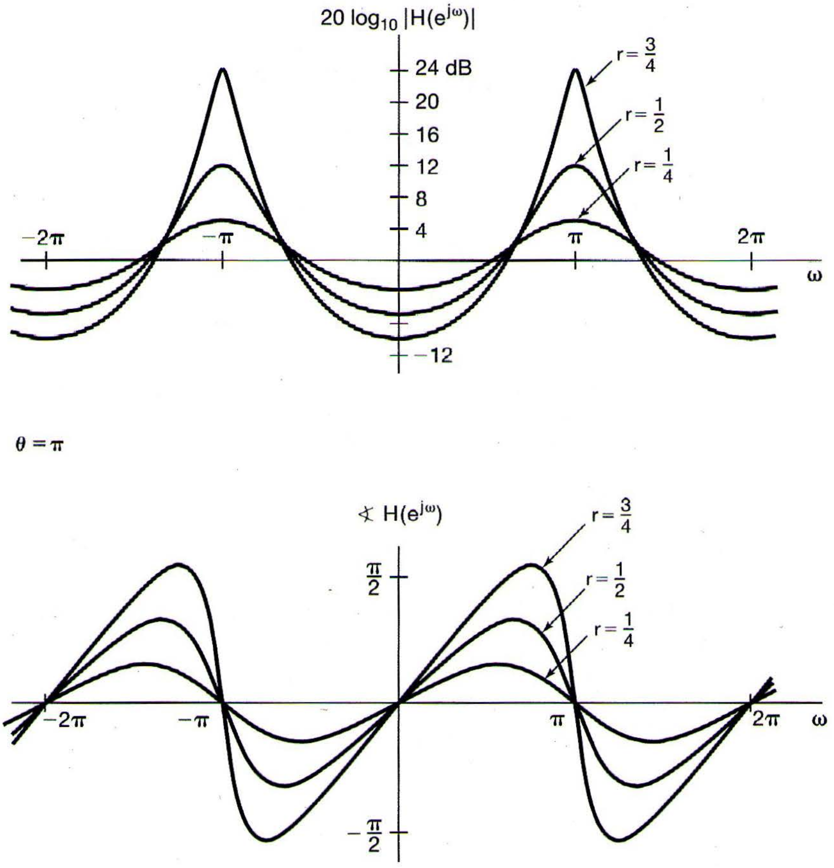
(d)

Figure 6.31 *Continued*

In this section, we have restricted attention to those causal first- and second-order systems that are stable and for which the frequency response can be defined. In particular, the causal system described by eq. (6.51) is unstable for $|a| \geq 1$. Also, the causal system described by eq. (6.56) is unstable if $r \geq 1$, and that described by eq. (6.71) is unstable if either $|d_1|$ or $|d_2|$ exceeds 1.

6.7 EXAMPLES OF TIME- AND FREQUENCY-DOMAIN ANALYSIS OF SYSTEMS

Throughout this chapter, we have illustrated the importance of viewing systems in both the time domain and the frequency domain and the importance of being aware of trade-offs in the behavior between the two domains. In this section, we illustrate some of these issues further. In Section 6.7.1, we discuss these trade-offs for continuous time in the context of an automobile suspension system. In Section 6.7.2, we discuss an important class of discrete-time filters referred to as moving-average or nonrecursive systems.



(e)

Figure 6.31 *Continued*

6.7.1 Analysis of an Automobile Suspension System

A number of the points that we have made concerning the characteristics and trade-offs in continuous-time systems can be illustrated in the interpretation of an automobile suspension system as a lowpass filter. Figure 6.32 shows a diagrammatic representation of a simple suspension system comprised of a spring and dashpot (shock absorber). The road surface can be thought of as a superposition of rapid small-amplitude changes in elevation (high frequencies), representing the roughness of the road surface, and gradual changes in elevation (low frequencies) due to the general topography. The automobile suspension system is generally intended to filter out rapid variations in the ride caused by the road surface (i.e., the system acts as a lowpass filter).

The basic purpose of the suspension system is to provide a smooth ride, and there is no sharp, natural division between the frequencies to be passed and those to be rejected. Thus, it is reasonable to accept and, in fact, prefer a lowpass filter that has a gradual

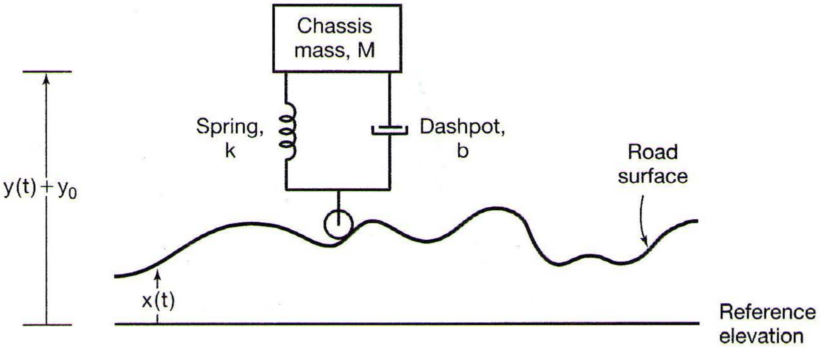


Figure 6.32 Diagrammatic representation of an automotive suspension system. Here, y_0 represents the distance between the chassis and the road surface when the automobile is at rest, $y(t) + y_0$ the position of the chassis above the reference elevation, and $x(t)$ the elevation of the road above the reference elevation.

transition from passband to stopband. Furthermore, the time-domain characteristics of the system are important. If the impulse response or step response of the suspension system exhibits ringing, then a large bump in the road (modeled as an impulse input) or a curb (modeled as a step input) will result in an uncomfortable oscillatory response. In fact, a common test for a suspension system is to introduce an excitation by depressing and then releasing the chassis. If the response exhibits ringing, it is an indication that the shock absorbers need to be replaced.

Cost and ease of implementation also play an important role in the design of automobile suspension systems. Many studies have been carried out to determine the most desirable frequency-response characteristics for suspension systems from the point of view of passenger comfort. In situations where the cost may be warranted, such as for passenger railway cars, intricate and costly suspension systems are used. For the automotive industry, cost is an important factor, and simple, less costly suspension systems are generally used. A typical automotive suspension system consists simply of the chassis connected to the wheels through a spring and a dashpot.

In the diagrammatic representation in Figure 6.32, y_0 represents the distance between the chassis and the road surface when the automobile is at rest, $y(t) + y_0$ the position of the chassis above the reference elevation, and $x(t)$ the elevation of the road above the reference elevation. The differential equation governing the motion of the chassis is then

$$M\frac{d^2y(t)}{dt^2} + b\frac{dy(t)}{dt} + ky(t) = kx(t) + b\frac{dx(t)}{dt}, \tag{6.76}$$

where M is the mass of the chassis and k and b are the spring and shock absorber constants, respectively. The frequency response of the system is

$$H(j\omega) = \frac{k + bj\omega}{(j\omega)^2M + b(j\omega) + k},$$

or

$$H(j\omega) = \frac{\omega_n^2 + 2\zeta\omega_n(j\omega)}{(j\omega)^2 + 2\zeta\omega_n(j\omega) + \omega_n^2}, \tag{6.77}$$

where

$$\omega_n = \sqrt{\frac{k}{M}} \quad \text{and} \quad 2\zeta\omega_n = \frac{b}{M}.$$

As in Section 6.5.2, the parameter ω_n is referred to as the undamped natural frequency and ζ as the damping ratio. A Bode plot of the log magnitude of the frequency response in eq. (6.77) can be constructed by using first-order and second-order Bode plots. The Bode plot for eq. (6.77) is sketched in Figure 6.33 for several different values of the damping ratio. Figure 6.34 illustrates the step response for several different values of the damping ratio.

As we saw in Section 6.5.2, the filter cutoff frequency is controlled primarily through ω_n, or equivalently for a chassis with a fixed mass, by an appropriate choice of spring constant k. For a given ω_n, the damping ratio is then adjusted through the damping factor b associated with the shock asorbers. As the natural frequency ω_n is decreased, the suspension will tend to filter out slower road variations, thus providing a smoother ride. On the other hand, we see from Figure 6.34 that the rise time of the system increases, and thus the system will feel more sluggish. On the one hand, it would be desirable to keep ω_n small to improve the lowpass filtering; on the other hand, it would be desirable to have ω_n large for a rapid time response. These, of course, are conflicting requirements and illustrate the need for a trade-off between time-domain and frequency-domain characteristics. Typically, a suspension system with a low value of ω_n, so that the rise time is long, is characterized as "soft" and one with a high value of ω_n, so that the rise time is short, is characterized as "hard." From Figures 6.33 and 6.34, we observe also that, as the damping ratio decreases, the frequency response of the system cuts off more sharply, but the overshoot and ringing in the step response tend to increase, another trade-off between the time and frequency

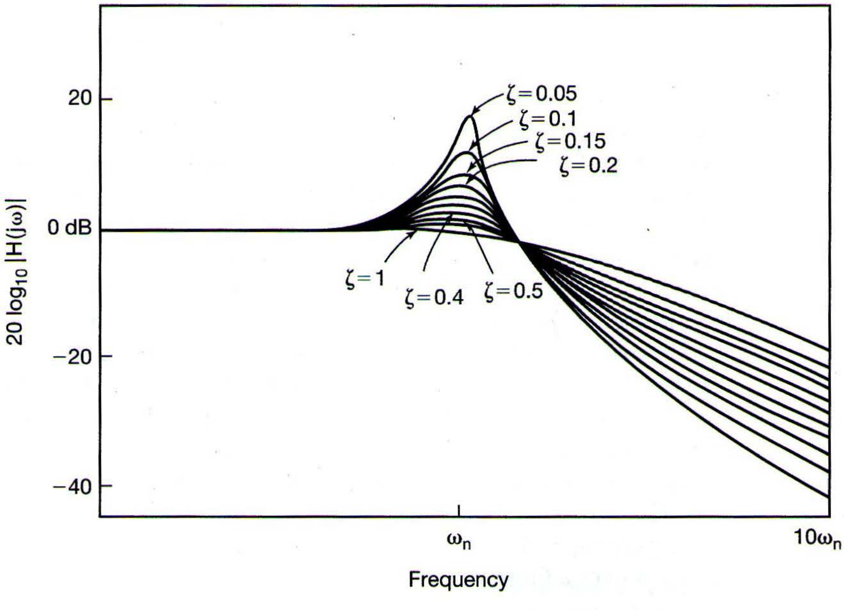


Figure 6.33 Bode plot for the magnitude of the frequency response of the automobile suspension system for several values of the damping ratio.

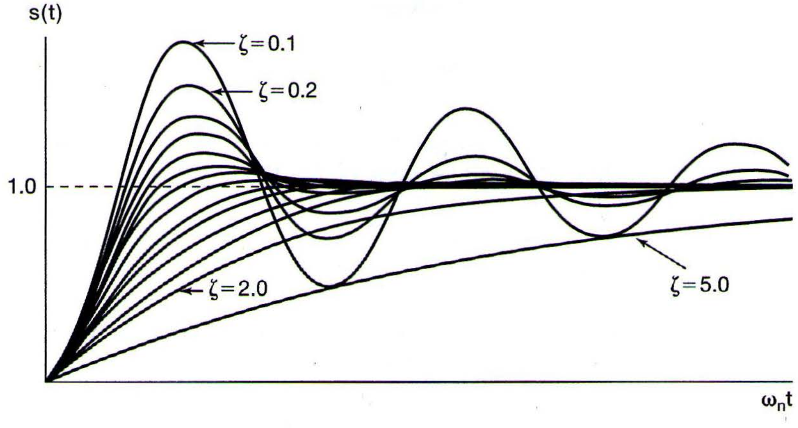


Figure 6.34 Step response of the automotive suspension system for various values of the damping ratio (ζ = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.2, 1.5, 2.0, 5.0).

domains. Generally, the shock absorber damping is chosen to have a rapid rise time and yet avoid overshoot and ringing. This choice corresponds to the critically damped case, with $\zeta = 1.0$, considered in Section 6.5.2.

6.7.2 Examples of Discrete-Time Nonrecursive Filters

In Section 3.11, we introduced the two basic classes of LTI filters described by difference equations, namely, recursive or infinite impulse response (IIR) filters and nonrecursive or finite impulse response (FIR) filters. Both of these classes of filters are of considerable importance in practice and have their own advantages and disadvantages. For example, recursive filters implemented as interconnections of the first- and second-order systems described in Section 6.6 provide a flexible class of filters that can be easily and efficiently implemented and whose characteristics can be adjusted by varying the number and the parameters of each of the component first- and second-order subsystems. On the other hand, as shown in Problem 6.64, it is not possible to design a causal, recursive filter with exactly linear phase, a property that we have seen is often desirable since, in that case, the effect of the phase on the output signal is a simple time delay. In contrast, as we show in this section, nonrecursive filters *can* have exactly linear phase. However, it is generally true that the same filter specifications require a higher order equation and hence more coefficients and delays when implemented with a nonrecursive equation, compared with a recursive difference equation. Consequently, for FIR filters, one of the principal trade-offs between the time and frequency domains is that increasing the flexibility in specifying the frequency domain characteristics of the filter, including, for example, achieving a higher degree of frequency selectivity, requires an FIR filter with an impulse response of longer duration.

One of the most basic nonrecursive filters, introduced in Section 3.11.2, is the moving-average filter. For this class of filters, the output is the average of the values of the input over a finite window:

$$y[n] = \frac{1}{N+M+1} \sum_{k=-N}^{M} x[n-k]. \tag{6.78}$$

The corresponding impulse response is a rectangular pulse, and the frequency response is

$$H(e^{j\omega}) = \frac{1}{N+M+1} e^{j\omega[(N-M)/2]} \frac{\sin[\omega(M+N+1)/2]}{\sin(\omega/2)}. \tag{6.79}$$

In Figure 6.35, we show the log magnitude for $M+N+1 = 33$ and $M+N+1 = 65$. The main, center lobe of each of these frequency responses corresponds to the effective passband of the corresponding filter. Note that, as the impulse response increases in length, the width of the main lobe of the magnitude of the frequency response decreases. This provides another example of the trade-off between the time and frequency domains. Specifically, in order to have a narrower passband, the filter in eqs. (6.78) and (6.79) must have a longer impulse response. Since the length of the impulse response of an FIR filter has a direct impact on the complexity of its implementation, this implies a trade-off between frequency selectivity and the complexity of the filter, a topic of central concern in filter design.

Moving-average filters are commonly applied in economic analysis in order to attenuate the short-term fluctuations in a variety of economic indicators in relation to longer term trends. In Figure 6.36, we illustrate the use of a moving-average filter of the form of eq. (6.78) on the weekly Dow Jones stock market index for a 10-year period. The weekly Dow Jones index is shown in Figure 6.36(a). Figure 6.36(b) is a 51-day moving average (i.e., $N = M = 25$) applied to that index, and Figure 6.36(c) is a 201-day moving average (i.e., $N = M = 100$) applied to the index. Both moving averages are considered useful, with the 51-day average tracking cyclical (i.e., periodic) trends that occur during the course of the year and the 201-day average primarily emphasizing trends over a longer time frame.

The more general form of a discrete-time nonrecursive filter is

$$y[n] = \sum_{k=-N}^{M} b_k x[n-k], \tag{6.80}$$

so that the output of this filter can be thought of as a weighted average of $N + M + 1$ neighboring points. The simple moving-average filter in eq. (6.78) then corresponds to setting all of these weights to the same value, namely, $1/(N+M+1)$. However, by choosing these coefficients in other ways, we have considerable flexibility in adjusting the filter's frequency response.

There are, in fact, a variety of techniques available for choosing the coefficients in eq. (6.80) so as to meet certain specifications on the filter, such as sharpening the transition band as much as possible for a filter of a given length (i.e., for N+M+1 fixed). These procedures are discussed in detail in a number of texts,[3] and although we do not discuss the procedures here, it is worth emphasizing that they rely heavily on the basic concepts and tools developed in this book. To illustrate how adjustment of the coefficients can influence

---

[3] See, for example, R. W. Hamming, *Digital Filters*, 3rd ed. (Englewood Cliffs, NJ: Prentice-Hall, Inc., 1989); A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing* (Englewood Cliffs, NJ: Prentice-Hall, Inc., 1989); and L. R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing* (Englewood Cliffs, NJ: Prentice-Hall, Inc., 1975).

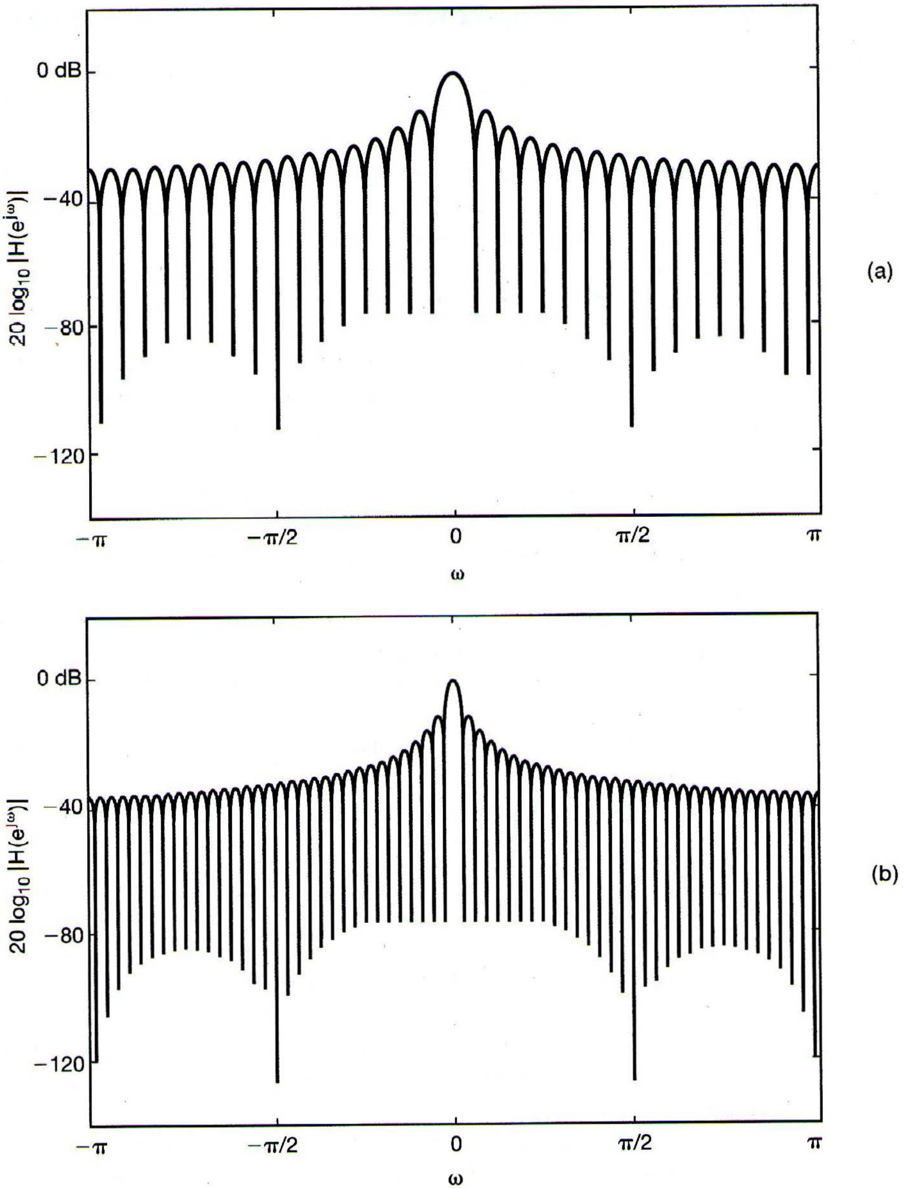


Figure 6.35 Log-magnitude plots for the moving-average filter of eqs. (6.78) and (6.79) for (a) $M + N + 1 = 33$ and (b) $M + N + 1 = 65$.

the response of the filter, let us consider a filter of the form of eq. (6.80), with $N = M = 16$ and the filter coefficients chosen to be

$$b_k = \begin{cases} \dfrac{\sin(2\pi k/33)}{\pi k}, & |k| \leq 32 \\ 0, & |k| > 32 \end{cases}. \tag{6.81}$$

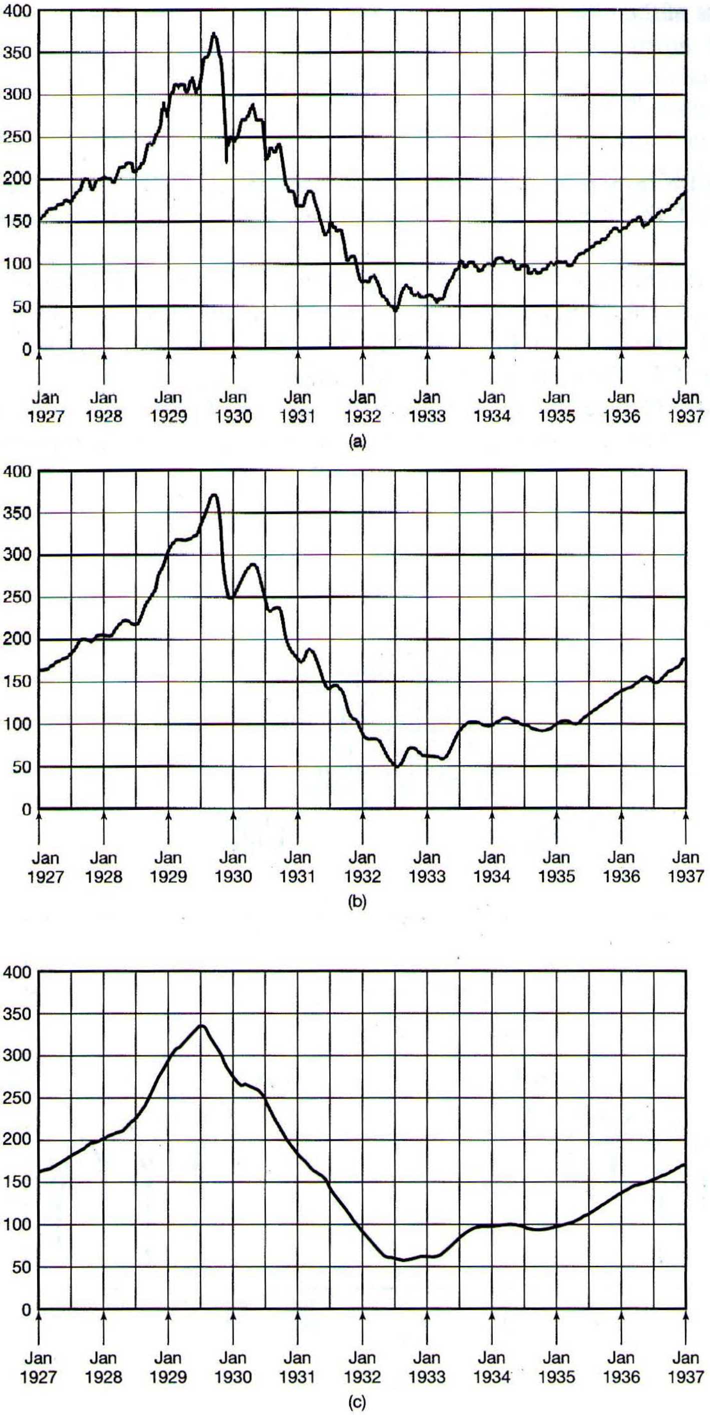


Figure 6.36 Effect of lowpass filtering on the Dow Jones weekly stock market index over a 10-year period using moving-average filters: (a) weekly index; (b) 51-day moving average applied to (a); (c) 201-day moving average applied to (a). The weekly stock market index and the two moving averages are discrete-time sequences. For clarity in the graphical display, the three sequences are shown here with their individual values connected by straight lines to form a continuous curve.

The impulse response of the filter is

$$h[n] = \begin{cases} \dfrac{\sin(2\pi n/33)}{\pi n}, & |n| \leq 32 \\ 0, & |n| > 32 \end{cases}. \tag{6.82}$$

Comparing this impulse response with eq. (6.20), we see that eq. (6.82) corresponds to truncating, for $|n| > 32$, the impulse response for the ideal lowpass filter with cutoff frequency $\omega_c = 2\pi/33$.

In general, the coefficients b_k can be adjusted so that the cutoff is at a desired frequency. For the example shown in Figure 6.37, the cutoff frequency was chosen to match approximately the cutoff frequency of Figure 6.35 for $N = M = 16$. Figure 6.37(a) shows the impulse response of the filter, and Figure 6.37(b) shows the log magnitude of the frequency response in dB. Comparing this frequency response to Figure 6.35, we observe that the passband of the filter has approximately the same width, but that the transition to

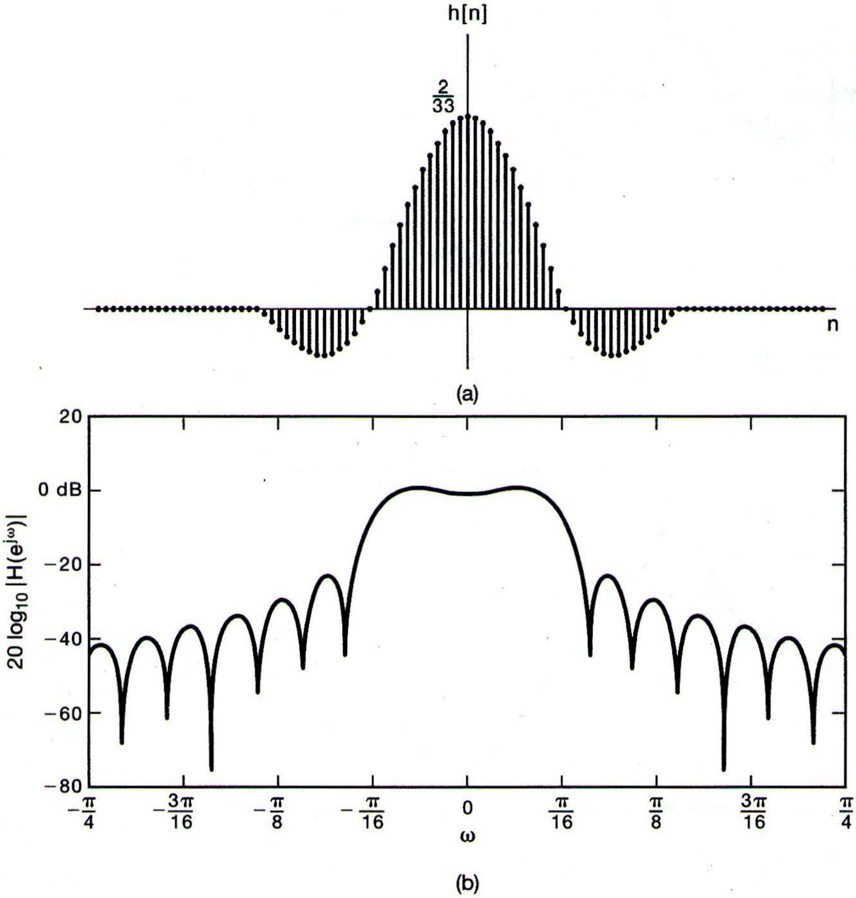


Figure 6.37 (a) Impulse response for the nonrecursive filter of eq. (6.82); (b) log magnitude of the frequency response of the filter.

the stopband is sharper. In Figures 6.38(a) and (b), the magnitudes (on a linear amplitude scale) of the two filters are shown for comparison. It should be clear from the comparison of the two examples that, by the intelligent choice of the weighting coefficients, the transition band can be sharpened. An example of a higher order lowpass filter ($N = M = 125$), with the coefficients determined through a numerical algorithm referred to as the Parks-McClellan algorithm,[4] is shown in Figure 6.39. This again illustrates the trade-off between the time and frequency domains: If we increase the length $N + M + 1$ of a filter, then, by a judicious choice of the filter coefficients in eq. (6.80), we can achieve sharper transition band behavior and a greater degree of frequency selectivity.

An important property of the examples we have given is that they all have zero or linear phase characteristics. For example, the phase of the moving-average filter of eq. (6.79) is $\omega[(N - M)/2]$. Also, since the impulse response in eq. (6.82) is real and even, the impulse response of the filter described by that equation is real and even, and thus has zero phase. From the symmetry properties of the Fourier transform of real signals, we know that *any* nonrecursive filter with an impulse response that is real and even will have a frequency response $H(e^{j\omega})$ that is real and even and, consequently, has zero phase. Such a filter, of course, is noncausal, since its impulse response $h[n]$ has nonzero values for $n < 0$. However, if a causal filter is required, then a simple change in the impulse response can achieve this, resulting in a system with *linear* phase. Specifically, since $h[n]$ is the impulse response of an FIR filter, it is identically zero outside a range of values centered at the origin

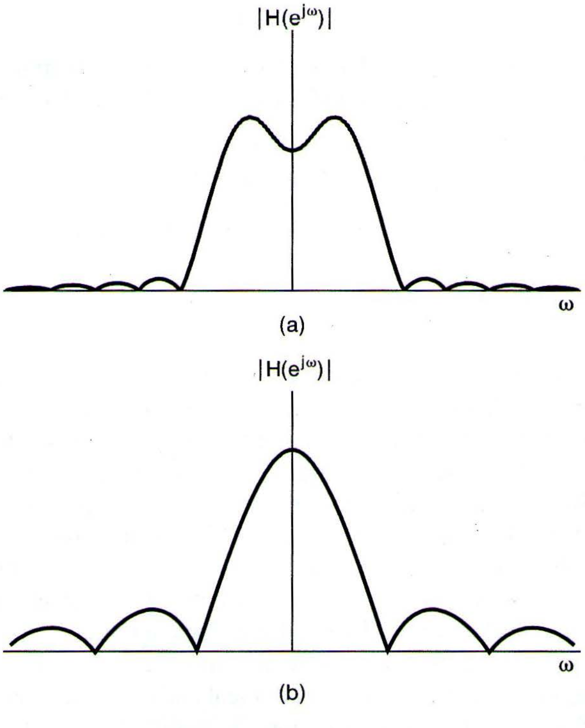


Figure 6.38 Comparison, on a linear amplitude scale, of the frequency responses of (a) Figure 6.37 and (b) Figure 6.35.

---

[4] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing* (Englewood Cliffs, NJ: Prentice-Hall, Inc., 1989), Chap. 7.

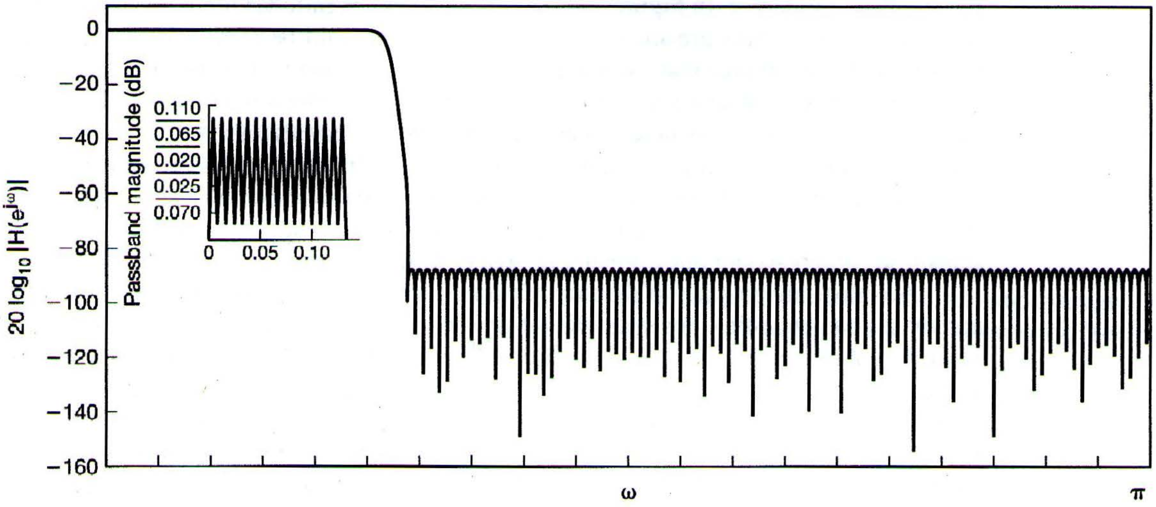


Figure 6.39 Lowpass nonrecursive filter with 251 coefficients designed to obtain the sharpest possible cutoff.

(i.e., $h[n] = 0$ for all $|n| > N$). If we now define the nonrecursive LTI system resulting from a simple N-step delay of $h[n]$, i.e.,

$$h_1[n] = h[n-N], \tag{6.83}$$

then $h_1[n] = 0$ for all $n < 0$, so that this LTI system is causal. Furthermore, from the time-shift property for discrete-time Fourier transforms, we see that the frequency response of the system is

$$H_1(e^{j\omega}) = H(e^{j\omega})e^{-j\omega N}. \tag{6.84}$$

Since $H(e^{j\omega})$ has zero phase, $H_1(e^{j\omega})$ does indeed have linear phase.

6.8 SUMMARY

In this chapter, we have built on the foundation of Fourier analysis of signals and systems developed in Chapters 3–5 in order to examine in more detail the characteristics of LTI systems and the effects they have on signals. In particular, we have taken a careful look at the magnitude and phase characteristics of signals and systems, and we have introduced log-magnitude and Bode plots for LTI systems. We have also discussed the impact of phase and phase distortion on signals and systems. This examination led us to understand the special role played by linear phase characteristics, which impart a constant delay at all frequencies and which, in turn, led to the concept of nonconstant group delay and dispersion associated with systems having nonlinear phase characteristics. Using these tools and insights, we took another look at frequency-selective filters and the time-frequency tradeoffs involved. We examined the properties of both ideal and non-ideal frequency-selective filters and saw that time-frequency considerations, causality constraints, and implementation issues frequently make non-ideal filters, with transition bands and tolerance limits in the passbands and stopbands, the preferred choice.

We also examined in detail the time-frequency characteristics of first- and second-order systems in both continuous and discrete time. We noted in particular the trade-off between the response time of these systems and the frequency-domain bandwidth. Since first- and second-order systems are the building blocks for more complex, higher order LTI systems, the insights developed for those basic systems are of considerable use in practice.

Finally, we presented several examples of LTI systems in order to illustrate many of the points developed in the chapter. In particular, we examined a simple model of an automobile suspension system to provide a concrete example of the time-response–frequency-response concerns that drive system design in practice. We also considered several examples of discrete-time nonrecursive filters, ranging from simple moving-average filters to higher order FIR filters designed to have enhanced frequency selectivity. We saw, in addition, that FIR filters can be designed so as to have exactly linear phase. These examples, the development of the tools of Fourier analysis that preceded them, and the insights those tools provide illustrate the considerable value of the methods of Fourier analysis in analyzing and designing LTI systems.

Chapter 6 Problems

The first section of problems belongs to the basic category, and the answers are provided in the back of the book. The remaining two sections contain problems belonging to the basic and advanced categories, respectively.

BASIC PROBLEMS WITH ANSWERS

6.1. Consider a continuous-time LTI system with frequency response $H(j\omega) = |H(j\omega)|\, e^{j\sphericalangle H(j\omega)}$ and real impulse response $h(t)$. Suppose that we apply an input $x(t) = \cos(\omega_0 t + \phi_0)$ to this system. The resulting output can be shown to be of the form

$$y(t) = Ax(t - t_0),$$

where A is a nonnegative real number representing an *amplitude-scaling* factor and t_0 is a time delay.

(a) Express A in terms of $|H(j\omega_0)|$.

(b) Express t_0 in terms of $\sphericalangle H(j\omega_0)$.

6.2. Consider a discrete-time LTI system with frequency response $H(e^{j\omega}) = |H(e^{j\omega})|e^{j\sphericalangle H(e^{j\omega})}$ and real impulse response $h[n]$. Suppose that we apply the input $x[n] = \sin(\omega_0 n + \phi_0)$ to this system. The resulting output can be shown to be of the form

$$y[n] = |H(e^{j\omega_0})|x[n - n_0],$$

provided that $\sphericalangle H(e^{j\omega_0})$ and ω_0 are related in a particular way. Determine this relationship.

6.3. Consider the following frequency response for a causal and stable LTI system:

$$H(j\omega) = \frac{1 - j\omega}{1 + j\omega}.$$

(a) Show that $|H(j\omega)| = A$, and determine the value of A.

(b) Determine which of the following statements is true about $\tau(\omega)$, the group delay of the system. (*Note:* $\tau(\omega) = -d(\sphericalangle H(j\omega))/d\omega$, where $\sphericalangle H(j\omega)$ is expressed in a form that does not contain any discontinuities.)

1. $\tau(\omega) = 0$ for $\omega > 0$
2. $\tau(\omega) > 0$ for $\omega > 0$
3. $\tau(\omega) < 0$ for $\omega > 0$

6.4. Consider a linear-phase discrete-time LTI system with frequency response $H(e^{j\omega})$ and real impulse response $h[n]$. The group delay function for such a system is defined as

$$\tau(\omega) = -\frac{d}{d\omega}\sphericalangle H(e^{j\omega}),$$

where $\sphericalangle H(e^{j\omega})$ has no discontinuities. Suppose that, for this system,

$$|H(e^{j\pi/2})| = 2,\ \ \sphericalangle H(e^{j0}) = 0,\ \text{ and }\ \tau\left(\frac{\pi}{2}\right) = 2.$$

Determine the output of the system for each of the following inputs:

(a) $\cos(\frac{\pi}{2}n)$ **(b)** $\sin(\frac{7\pi}{2}n + \frac{\pi}{4})$

6.5. Consider a continuous-time ideal bandpass filter whose frequency response is

$$H(j\omega) = \begin{cases} 1, & \omega_c \le |\omega| \le 3\omega_c \\ 0, & \text{elsewhere} \end{cases}.$$

(a) If $h(t)$ is the impulse response of this filter, determine a function $g(t)$ such that

$$h(t) = \left(\frac{\sin \omega_c t}{\pi t}\right) g(t).$$

(b) As ω_c is increased, does the impulse response of the filter get more concentrated or less concentrated about the origin?

6.6. Consider a discrete-time ideal highpass filter whose frequency response is specified as

$$H(e^{j\omega}) = \begin{cases} 1, & \pi - \omega_c \le |\omega| \le \pi \\ 0, & |\omega| < \pi - \omega_c \end{cases}$$

(a) If $h[n]$ is the impulse response of this filter, determine a function $g[n]$ such that

$$h[n] = \left(\frac{\sin \omega_c n}{\pi n}\right) g[n].$$

(b) As ω_c is increased, does the impulse response of the filter get more concentrated or less concentrated about the origin?

6.7. A continuous-time lowpass filter has been designed with a passband frequency of 1,000 Hz, a stopband frequency of 1,200 Hz, passband ripple of 0.1, and stopband ripple of 0.05. Let the impulse response of this lowpass filter be denoted by $h(t)$. We wish to convert the filter into a bandpass filter with impulse response

$$g(t) = 2h(t)\cos(4{,}000\pi t).$$

Assuming that $|H(j\omega)|$ is negligible for $|\omega| > 4{,}000\pi$, answer the following questions:

(a) If the passband ripple for the bandpass filter is constrained to be 0.1, what are the two passband frequencies associated with the bandpass filter?

(b) If the stopband ripple for the bandpass filter is constrained to be 0.05, what are the two stopband frequencies associated with the bandpass filter?

6.8. A causal, nonideal lowpass filter is designed with frequency response $H(e^{j\omega})$. The difference equation relating the input $x[n]$ and output $y[n]$ for this filter is specified as

$$y[n] = \sum_{k=1}^{N} a_k y[n-k] + \sum_{k=0}^{M} b_k x[n-k].$$

The filter also satisfies the following specifications for the magnitude of its frequency response:

$$\begin{aligned} \text{passband frequency} &= \omega_p, \\ \text{passband tolerance} &= \delta_p, \\ \text{stopband frequency} &= \omega_s, \\ \text{stopband tolerance} &= \delta_s. \end{aligned}$$

Now consider a causal LTI system whose input and output are related by the difference equation

$$y[n] = \sum_{k=1}^{N} (-1)^k a_k y[n-k] + \sum_{k=0}^{M} (-1)^k b_k x[n-k].$$

Show that this filter has a passband with a tolerance of δ_p, and specify the corresponding location of the passband.

6.9. Consider a continuous-time causal and stable LTI system whose input $x(t)$ and output $y(t)$ are related by the differential equation

$$\frac{dy(t)}{dt} + 5y(t) = 2x(t).$$

What is the final value $s(\infty)$ of the step response $s(t)$ of this filter? Also, determine the value of t_0 for which

$$s(t_0) = s(\infty)\left[1 - \frac{1}{e^2}\right].$$

6.10. For each first-order system whose frequency response is as follows, specify the straight-line approximation of the Bode magnitude plot:

(a) $40\left(\frac{j\omega+0.1}{j\omega+40}\right)$ **(b)** $0.04\left(\frac{j\omega+50}{j\omega+0.2}\right)$

6.11. For each second-order system whose frequency response is as follows, specify the straight-line approximation of the Bode magnitude plot:

(a) $\frac{250}{(j\omega)^2+50.5j\omega+25}$ **(b)** $0.02\frac{j\omega+50}{(j\omega)^2+0.2j\omega+1}$

6.12. A continuous-time LTI system S with frequency response $H(j\omega)$ is constructed by cascading two continuous-time LTI systems with frequency responses $H_1(j\omega)$ and $H_2(j\omega)$, respectively. Figures P6.12(a) and P6.12(b) show the straight-line approximations of the Bode magnitude plots of $H_1(j\omega)$ and $H(j\omega)$, respectively. Specify $H_2(j\omega)$.

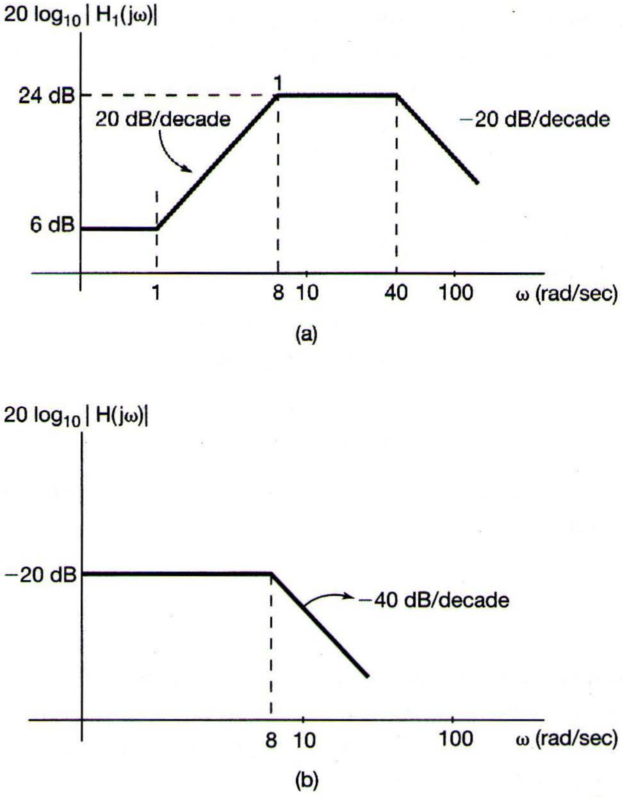


Figure P6.12

6.13. The straight-line approximation of the Bode magnitude plot of a second-order continuous-time LTI system S is shown in Figure P6.13. S may be constructed by

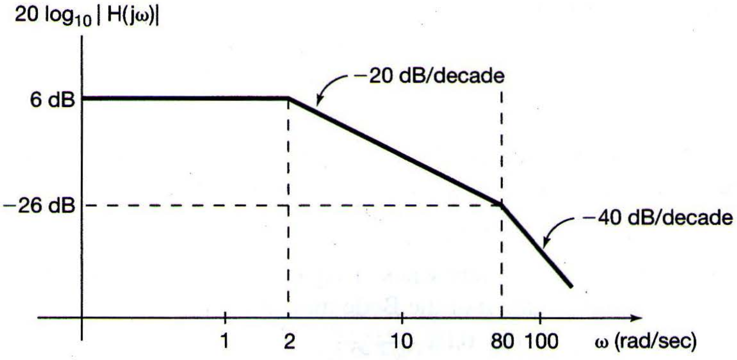


Figure P6.13

either connecting two first-order systems S_1 and S_2 in cascade or two first-order systems S_3 and S_4 in parallel. Determine which, if any, of the following statements are true or false. Justify your answers.

(a) The frequency responses of S_1 and S_2 may be determined uniquely.

(b) The frequency responses of S_3 and S_4 may be determined uniquely.

6.14. The straight-line approximation of the Bode magnitude plot of a causal and stable continuous-time LTI system S is shown in Figure P6.14. Specify the frequency response of a system that is the inverse of S.

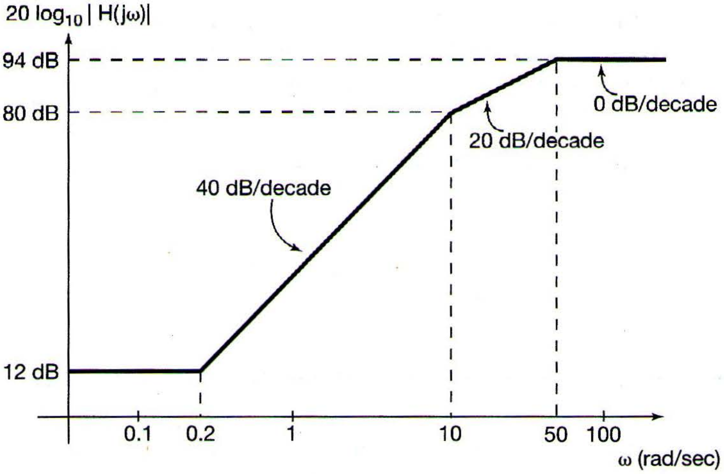


Figure P6.14

6.15. For each of the following second-order differential equations for causal and stable LTI systems, determine whether the corresponding impulse response is underdamped, overdamped, or critically damped:

(a) $\frac{d^2y(t)}{dt^2} + 4\frac{dy(t)}{dt} + 4y(t) = x(t)$

(b) $5\frac{d^2y(t)}{dt^2} + 4\frac{dy(t)}{dt} + 5y(t) = 7x(t)$

(c) $\frac{d^2y(t)}{dt^2} + 20\frac{dy(t)}{dt} + y(t) = x(t)$

(d) $5\frac{d^2y(t)}{dt^2} + 4\frac{dy(t)}{dt} + 5y(t) = 7x(t) + \frac{1}{3}\frac{dx(t)}{dt}$

6.16. A particular first-order causal and stable discrete-time LTI system has a step response whose maximum overshoot is 50% of its final value. If the final value is 1, determine a difference equation relating the input $x[n]$ and output $y[n]$ of this filter.

6.17. For each of the following second-order difference equations for causal and stable LTI systems, determine whether or not the step response of the system is oscillatory:

(a) $y[n] + y[n-1] + \frac{1}{4}y[n-2] = x[n]$

(b) $y[n] - y[n-1] + \frac{1}{4}y[n-2] = x[n]$

6.18. Consider the continuous-time LTI system implemented as the RC circuit shown in Figure P6.18. The voltage source $x(t)$ is considered the input to this system. The voltage $y(t)$ across the capacitor is considered the system output. Is it possible for the step response of the system to have an oscillatory behavior?

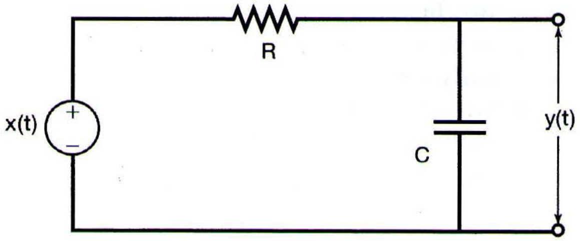


Figure P6.18

6.19. Consider the continuous-time LTI system implemented as the *RLC* circuit shown in Figure P6.19. The voltage source $x(t)$ is considered the input to this system. The voltage $y(t)$ across the capacitor is considered the system output. How should R, L, and C be related so that there is no oscillation in the step response?

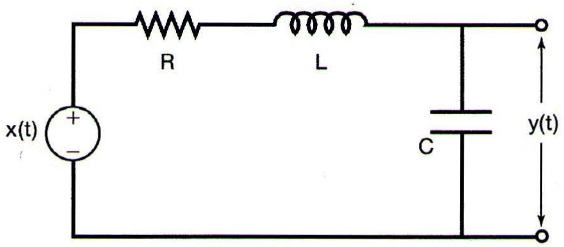


Figure P6.19

6.20. Consider a nonrecursive filter with the impulse response shown in Figure P6.20. What is the group delay as a function of frequency for this filter?

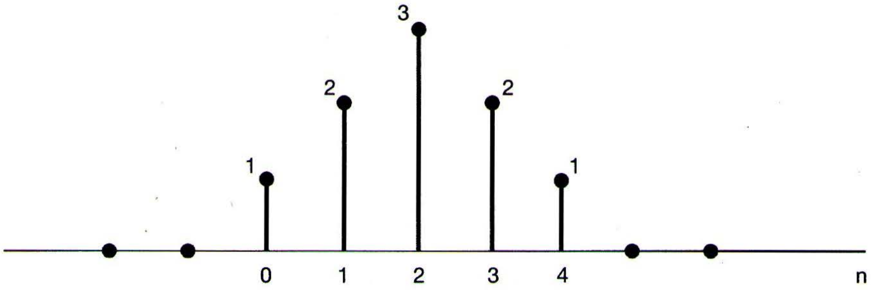


Figure P6.20

BASIC PROBLEMS

6.21. A causal LTI filter has the frequency response $H(j\omega)$ shown in Figure P6.21. For each of the input signals given below, determine the filtered output signal $y(t)$.

(a) $x(t) = e^{jt}$ **(b)** $x(t) = (\sin \omega_0 t)u(t)$

(c) $X(j\omega) = \frac{1}{(j\omega)(6+j\omega)}$ **(d)** $X(j\omega) = \frac{1}{2+j\omega}$

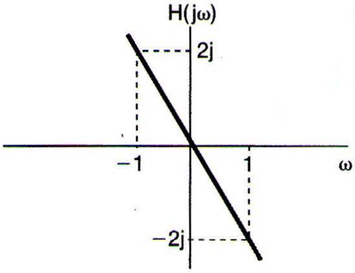


Figure P6.21

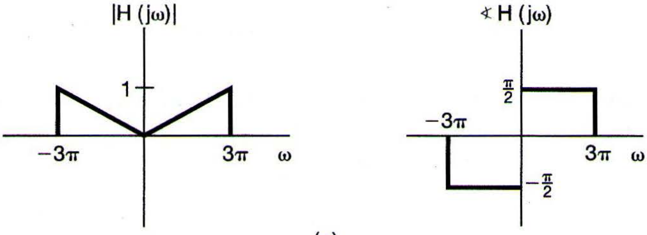
6.22. Shown in Figure P6.22(a) is the frequency response $H(j\omega)$ of a continuous-time filter referred to as a lowpass differentiator. For each of the input signals $x(t)$ below, determine the filtered output signal $y(t)$.

(a) $x(t) = \cos(2\pi t + \theta)$

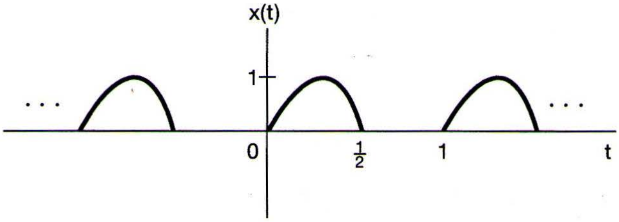
(b) $x(t) = \cos(4\pi t + \theta)$

(c) $x(t)$ is a half-wave rectified sine wave of period, as sketched in Figure P6.22(b).

$$x(t) = \begin{cases} \sin 2\pi t, & m \le t \le (m + \frac{1}{2}) \\ 0, & (m + \frac{1}{2}) \le t \le m \text{ for any integer } m \end{cases}$$



(a)



(b)

Figure P6.22

6.23. Shown in Figure P6.23 is $|H(j\omega)|$ for a lowpass filter. Determine and sketch the impulse response of the filter for each of the following phase characteristics:

(a) $\sphericalangle H(j\omega) = 0$

(b) $\sphericalangle H(j\omega) = \omega T$, where T is a constant

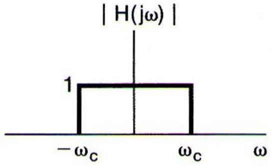


Figure P6.23

(c) $\sphericalangle H(j\omega) = \begin{cases} \frac{\pi}{2}, & \omega > 0 \\ \frac{-\pi}{2}, & \omega < 0 \end{cases}$

6.24. Consider a continuous-time lowpass filter whose impulse response $h(t)$ is known to be real and whose frequency response magnitude is given as:

$$|H(j\omega)| = \begin{cases} 1, & |\omega| \leq 200\pi \\ 0, & \text{otherwise} \end{cases}.$$

(a) Determine and sketch the real-valued impulse response $h(t)$ for this filter when the corresponding group delay function is specified as:

(i) $\tau(\omega) = 5$ **(ii)** $\tau(\omega) = \frac{5}{2}$ **(iii)** $\tau(\omega) = -\frac{5}{2}$

(b) If the impulse response $h(t)$ had not been specified to be real, would knowledge of $|H(j\omega)|$ and $\tau(\omega)$ be sufficient to determine $h(t)$ uniquely? Justify your answer.

6.25. By computing the group delay at two selected frequencies, verify that each of the following frequency responses has nonlinear phase.

(a) $H(j\omega) = \frac{1}{j\omega+1}$ **(b)** $H(j\omega) = \frac{1}{(j\omega+1)^2}$ **(c)** $H(j\omega) = \frac{1}{(j\omega+1)(j\omega+2)}$

6.26. Consider an ideal highpass filter whose frequency response is specified as

$$H(j\omega) = \begin{cases} 1, & |\omega| > \omega_c \\ 0, & \text{otherwise} \end{cases}.$$

(a) Determine the impulse response $h(t)$ for this filter.

(b) As ω_c is increased, does $h(t)$ get more or less concentrated about the origin?

(c) Determine $s(0)$ and $s(\infty)$, where $s(t)$ is the step response of the filter.

6.27. The output $y(t)$ of a causal LTI system is related to the input $x(t)$ by the differential equation

$$\frac{dy(t)}{dt} + 2y(t) = x(t).$$

(a) Determine the frequency response

$$H(j\omega) = \frac{Y(j\omega)}{X(j\omega)}$$

of the system, and sketch its Bode plot.

(b) Specify, as a function of frequency, the group delay associated with this system.

(c) If $x(t) = e^{-t}u(t)$, determine $Y(j\omega)$, the Fourier transform of the output.

(d) Using the technique of partial-fraction expansion, determine the output $y(t)$ for the input $x(t)$ in part (c).

(e) Repeat parts (c) and (d), first if the input has as its Fourier transform

(i) $X(j\omega) = \frac{1+j\omega}{2+j\omega}$,

then if

(ii) $X(j\omega) = \frac{2+j\omega}{1+j\omega}$,

and finally, if

(iii) $X(j\omega) = \frac{1}{(2+j\omega)(1+j\omega)}$

6.28. **(a)** Sketch the Bode plots for the following frequency responses:

(i) $1 + (j\omega/10)$ **(ii)** $1 - (j\omega/10)$

(iii) $\frac{16}{(j\omega+2)^4}$ **(iv)** $\frac{1-(j\omega/10)}{1+j\omega}$

(v) $\frac{(j\omega/10)-1}{1+j\omega}$ **(vi)** $\frac{1+(j\omega/10)}{1+j\omega}$

(vii) $\frac{1-(j\omega/10)}{(j\omega)^2+(j\omega)+1}$ **(viii)** $\frac{10+5j\omega+10(j\omega)^2}{1+(j\omega/10)}$

(ix) $1 + j\omega + (j\omega)^2$ **(x)** $1 - j\omega + (j\omega)^2$

(xi) $\frac{(j\omega+10)(10j\omega+1)}{[(j\omega/100+1)][((j\omega)^2+j\omega+1)]}$

(b) Determine and sketch the impulse response and the step response for the system with frequency response (iv). Do the same for the system with frequency response (vi).

The system given in (iv) is often referred to as a non-minimum-phase system, while the system specified in (vi) is referred to as being a minimum phase. The corresponding impulse responses of (iv) and (vi) are referred to as a non-minimum-phase signal and a minimum-phase signal, respectively. By comparing the Bode plots of these two frequency responses, we can see that they have identical magnitudes; however, the magnitude of the *phase* of the system of (iv) is larger than for the system of (vi).

We can also note differences in the time-domain behavior of the two systems. For example, the impulse response of the minimum-phase system has more of its energy concentrated near $t = 0$ than does the impulse response of the non-minimum-phase system. In addition, the step response of (iv) initially has the opposite sign from its asymptotic value as $t \to \infty$, while this is not the case for the system of (vi).

The important concept of minimum- and non-minimum-phase systems can be extended to more general LTI systems than the simple first-order systems we have treated here, and the distinguishing characteristics of these systems can be described far more thoroughly than we have done.

6.29. An LTI system is said to have *phase lead* at a particular frequency $\omega = \omega_0$ if $\sphericalangle H(j\omega_0) > 0$. The terminology stems from the fact that if $e^{j\omega_0 t}$ is the input to this system, then the phase of the output will exceed, or lead, the phase of the input. Similarly, if $\sphericalangle H(j\omega_0) < 0$, the system is said to have *phase lag* at this frequency. Note that the system with frequency response

$$\frac{1}{1 + j\omega\tau}$$

has phase lag for all $\omega > 0$, while the system with frequency response

$$1 + j\omega\tau$$

has phase lead for all $\omega > 0$.

(a) Construct the Bode plots for the following two systems. Which has phase lead and which phase lag? Also, which one amplifies signals at certain frequencies?

(i) $\frac{1+(j\omega/10)}{1+10j\omega}$ **(ii)** $\frac{1+10j\omega}{1+(j\omega/10)}$

(b) Repeat part (a) for the following three frequency responses:

(i) $\frac{(1+(j\omega/10))^2}{(1+10j\omega)^3}$ **(ii)** $\frac{1+j\omega/10}{100(j\omega)^2+10j\omega+1}$ **(iii)** $\frac{1+10j\omega}{0.01(j\omega)^2+0.2j\omega+1}$

6.30. Let $h(t)$ have the Bode plot depicted in Figure P6.30. The dashed lines in the figure represent straight-line approximations. Sketch the Bode plots for $10h(10t)$.

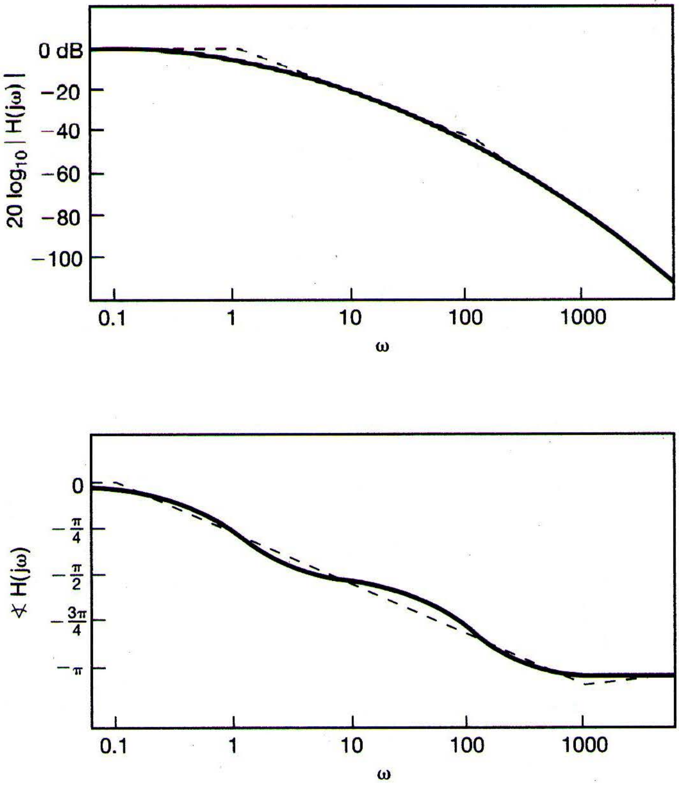


Figure P6.30

6.31. An integrator has as its frequency response

$$H(j\omega) = \frac{1}{j\omega} + \pi\,\delta(\omega),$$

where the impulse at $\omega = 0$ is a result of the fact that the integration of a constant input from $t = -\infty$ results in an infinite output. Thus, if we avoid inputs that are constant, or equivalently, only examine $H(j\omega)$ for $\omega > 0$, we see that

$$20 \log |H(j\omega)| = -20 \log(\omega),$$

$$\sphericalangle H(j\omega) = \frac{-\pi}{2}.$$

In other words, the Bode plot for an integrator, as illustrated in Figure P6.31, consists of two straight-line plots. These plots reflect the principal characteristics of an integrator: a phase shift of $-90°$ at all positive values of frequency and the amplification of low frequencies.

(a) A useful, simple model of an electric motor is an LTI system with input equal to the applied voltage and output given by the motor shaft angle. This system can be visualized as the cascade of a stable LTI system (with the voltage as input and shaft angular velocity as output) and an integrator (representing the integration of the angular velocity). Often, a model of first-order system is used for the first part of the cascade. Assuming, for example, that this first-order system has a time constant of 0.1 second, we obtain an overall motor frequency response of

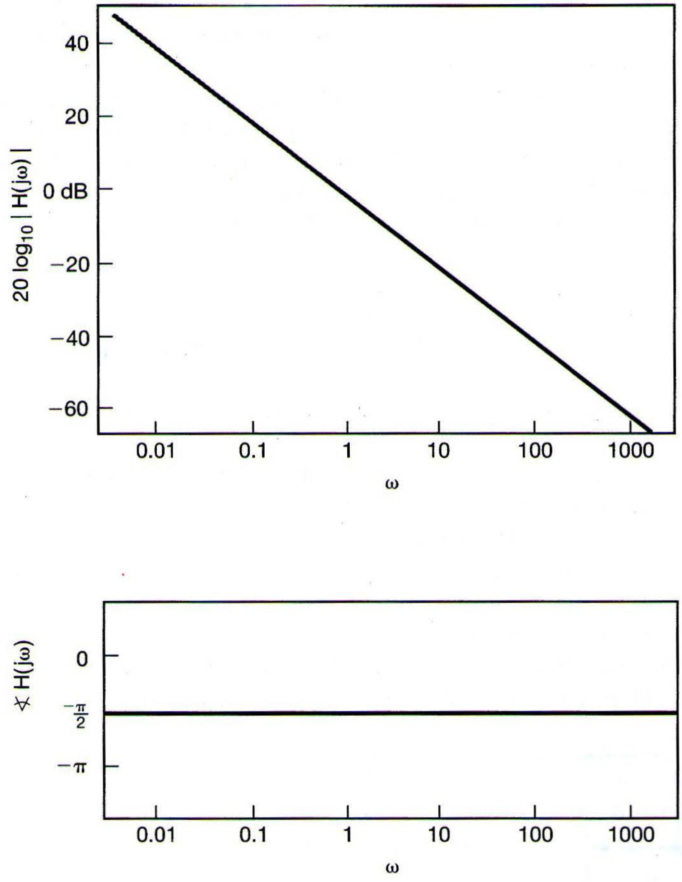


Figure P6.31

the form

$$H(j\omega) = \frac{1}{j\omega(1 + j\omega/10)} + \pi\,\delta(\omega).$$

Sketch the Bode plot for the system for $\omega > 0.001$.

(b) Sketch the Bode plot for a differentiator.

(c) Do the same for systems with the following frequency responses:

(i) $H(j\omega) = \frac{j\omega}{1+j\omega/100}$

(ii) $H(j\omega) = \frac{j\omega}{(1+(j\omega)/10+(j\omega)^2/100)}$

6.32. Consider the system depicted in Figure P6.32. This "compensator" box is a continuous-time LTI system.

(a) Suppose that it is desired to choose the frequency response of the compensator so that the overall frequency response $H(j\omega)$ of the cascade satisfies the following two conditions:

1. The log magnitude of $H(j\omega)$ has a slope of -40 dB/decade beyond $\omega = 1{,}000$.
2. For $0 < \omega < 1{,}000$, the log magnitude of $H(j\omega)$ should be between -10 dB and 10 dB.

Design a suitable compensator (that is, determine a frequency response for a compensator that meets the preceding requirements), and draw the Bode plot for the resulting $H(j\omega)$.

(b) Repeat (a) if the specifications on the log magnitude of $H(j\omega)$ are as follows:

1. It should have a slope of $+20$ dB/decade for $0 < \omega < 10$.
2. It should be between $+10$ and $+30$ dB for $10 < \omega < 100$.
3. It should have a slope of -20 dB/decade for $100 < \omega < 1{,}000$.
4. It should have a slope of -40 dB/decade for $\omega > 1{,}000$.

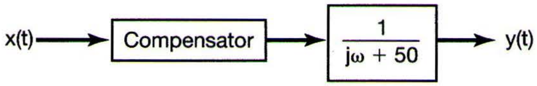


Figure P6.32

6.33. Figure P6.33 shows a system commonly used to obtain a highpass filter from a lowpass filter and vice versa.

(a) Show that, if $H(j\omega)$ is a lowpass filter with cutoff frequency ω_{lp}, the overall system corresponds to an ideal highpass filter. Determine the system's cutoff frequency and sketch its impulse response.

(b) Show that, if $H(j\omega)$ is an ideal highpass filter with cutoff frequency ω_{hp}, the overall system corresponds to an ideal lowpass filter, and determine the cutoff frequency of the system.

(c) If the interconnection of Figure P6.33 is applied to an ideal discrete-time lowpass filter, will the resulting system be an ideal discrete-time highpass filter?

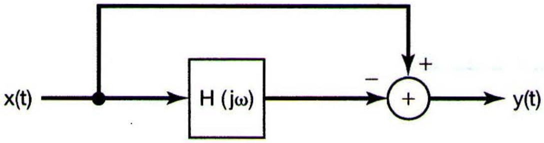


Figure P6.33

6.34. In Problem 6.33, we considered a system commonly used to obtain a highpass filter from a lowpass filter and vice versa. In this problem, we explore the system further and, in particular, consider a potential difficulty if the phase of $H(j\omega)$ is not properly chosen.

(a) Referring to Figure P6.33, let us assume that $H(j\omega)$ is real and as shown in Figure P6.34. Then

$$1 - \delta_1 < H(j\omega) < 1 + \delta_1, \quad 0 \leq \omega \leq \omega_1,$$

$$-\delta_2 < H(j\omega) < +\delta_2, \quad \omega_2 < \omega.$$

Determine and sketch the resulting frequency response of the overall system of Figure P6.33. Does the resulting system correspond to an approximation to a highpass filter?

(b) Now let $H(j\omega)$ in Figure P6.33 be of the form

$$H(j\omega) = H_1(j\omega)e^{j\theta(\omega)}, \qquad \text{(P6.34–1)}$$

where $H_1(j\omega)$ is identical to Figure P6.34 and $\theta(\omega)$ is an unspecified phase characteristic. With $H(j\omega)$ in this more general form, does it still correspond to an approximation to a lowpass filter?

(c) Without making any assumptions about $\theta(\omega)$, determine and sketch the tolerance limits on the magnitude of the frequency response of the overall system of Figure P6.33.

(d) If $H(j\omega)$ in Figure P6.33 is an approximation to a lowpass filter with unspecified phase characteristics, will the overall system in that figure necessarily correspond to an approximation to a highpass filter?

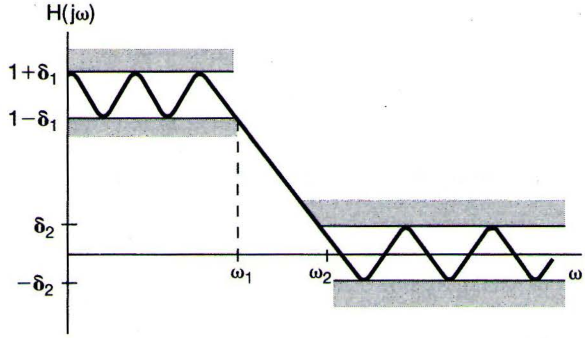


Figure P6.34

6.35. Shown in Figure P6.35 is the frequency response $H(e^{j\omega})$ of a discrete-time differentiator. Determine the output signal $y[n]$ as a function of ω_0 if the input $x[n]$ is

$$x[n] = \cos[\omega_0 n + \theta].$$

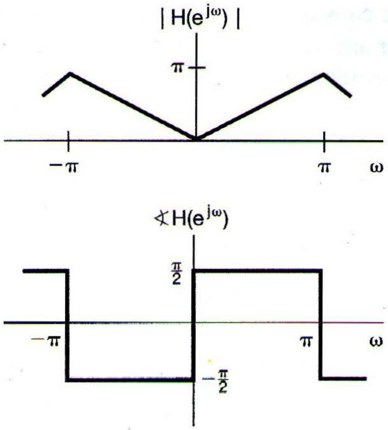


Figure P6.35

6.36. Consider a discrete-time lowpass filter whose impulse response $h[n]$ is known to be real and whose frequency response magnitude in the region $-\pi \leq \omega \leq \pi$ is given as:

$$|H(e^{j\omega})| = \begin{cases} 1, & |\omega| \leq \frac{\pi}{4} \\ 0, & \text{otherwise} \end{cases}.$$

Determine and sketch the real-valued impulse response $h[n]$ for this filter when the corresponding group delay function is specified as:
(a) $\tau(\omega) = 5$ **(b)** $\tau(\omega) = \frac{5}{2}$ **(c)** $\tau(\omega) = -\frac{5}{2}$

6.37. Consider a causal LTI system whose frequency response is given as:

$$H(e^{j\omega}) = e^{-j\omega}\frac{1 - \frac{1}{2}e^{j\omega}}{1 - \frac{1}{2}e^{-j\omega}}.$$

(a) Show that $|H(e^{j\omega})|$ is unity at all frequencies.
(b) Show that

$$\sphericalangle H(e^{j\omega}) = -\omega - 2\tan^{-1}\left(\frac{\frac{1}{2}\sin\omega}{1 - \frac{1}{2}\cos\omega}\right).$$

(c) Show that the group delay for this filter is given by

$$\tau(\omega) = \frac{\frac{3}{4}}{\frac{5}{4} - \cos\omega}.$$

Sketch $\tau(\omega)$.
(d) What is the output of this filter when the input is $\cos(\frac{\pi}{3}n)$?

6.38. Consider an ideal bandpass filter whose frequency response in the region $-\pi \leq \omega \leq \pi$ is specified as

$$H(e^{j\omega}) = \begin{cases} 1, & \frac{\pi}{2} - \omega_c \leq |\omega| \leq \frac{\pi}{2} + \omega_c \\ 0, & \text{otherwise} \end{cases}.$$

Determine and sketch the impulse response $h[n]$ for this filter when

(a) $\omega_c = \frac{\pi}{5}$

(b) $\omega_c = \frac{\pi}{4}$

(c) $\omega_c = \frac{\pi}{3}$

As ω_c is increased, does $h[n]$ get more or less concentrated about the origin?

6.39. Sketch the log magnitude and phase of each of the following frequency responses.

(a) $1 + \frac{1}{2}e^{-j\omega}$ **(b)** $1 + 2e^{-j\omega}$

(c) $1 - 2e^{-j\omega}$ **(d)** $1 + 2e^{-j2\omega}$

(e) $\frac{1}{(1+\frac{1}{2}e^{-j\omega})^3}$ **(f)** $\frac{1+\frac{1}{2}e^{-j\omega}}{1-\frac{1}{2}e^{-j\omega}}$

(g) $\frac{1+2e^{-j\omega}}{1+\frac{1}{2}e^{-j\omega}}$ **(h)** $\frac{1-2e^{-j\omega}}{1+\frac{1}{2}e^{-j\omega}}$

(i) $\frac{1}{(1-\frac{1}{4}e^{-j\omega})(1-\frac{3}{4}e^{-j\omega})}$ **(j)** $\frac{1}{(1-\frac{1}{4}e^{-j\omega})(1+\frac{3}{4}e^{-j\omega})}$

(k) $\frac{1+2e^{-2j\omega}}{(1-\frac{1}{2}e^{-j\omega})^2}$

6.40. Consider an ideal discrete-time lowpass filter with impulse response $h[n]$ and for which the frequency response $H(e^{j\omega})$ is that shown in Figure P6.40. Let us consider obtaining a new filter with impulse response $h_1[n]$ and frequency response $H_1(e^{j\omega})$ as follows:

$$h_1[n] = \begin{cases} h[n/2], & n \text{ even} \\ 0, & n \text{ odd} \end{cases}$$

This corresponds to inserting a sequence value of zero between each sequence value of $h[n]$. Determine and sketch $H_1(e^{j\omega})$ and state the class of ideal filters to which it belongs (e.g., lowpass, highpass, bandpass, multiband, etc.).

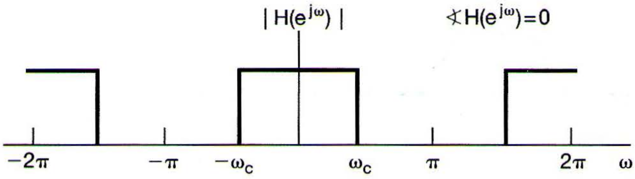


Figure P6.40

6.41. A particular causal LTI system is described by the difference equation

$$y[n] - \frac{\sqrt{2}}{2}y[n-1] + \frac{1}{4}y[n-2] = x[n] - x[n-1].$$

(a) Find the impulse response of this system.

(b) Sketch the log magnitude and the phase of the frequency response of the system.

6.42. **(a)** Consider two LTI systems with the following frequency responses:

$$H_1(e^{j\omega}) = \frac{1 + \frac{1}{2}e^{-j\omega}}{1 + \frac{1}{4}e^{-j\omega}},$$

$$H_2(e^{j\omega}) = \frac{\frac{1}{2} + e^{-j\omega}}{1 + \frac{1}{4}e^{-j\omega}}.$$

Show that both of these frequency responses have the same magnitude function [i.e., $|H_1(e^{j\omega})| = |H_2(e^{j\omega})|$], but the group delay of $H_2(e^{j\omega})$ is greater than the group delay of $H_1(e^{j\omega})$ for $\omega > 0$.

(b) Determine and sketch the impulse and step responses of the two systems.

(c) Show that

$$H_2(e^{j\omega}) = G(e^{j\omega})H_1(e^{j\omega}),$$

where $G(e^{j\omega})$ is an *all-pass system* [i.e., $|G(e^{j\omega})| = 1$ for all ω].

6.43. When designing filters with highpass or bandpass characteristics, it is often convenient first to design a lowpass filter with the desired passband and stopband specifications and then to transform this prototype filter to the desired highpass or bandpass filter. Such transformations are called lowpass-to-highpass or highpass-to-lowpass transformations. Designing filters in this manner is convenient because it requires us only to formulate our filter design algorithms for the class of filters with lowpass characteristics. As one example of such a procedure, consider a discrete-time lowpass filter with impulse response $h_{\mathrm{lp}}[n]$ and frequency response $H_{\mathrm{lp}}(e^{j\omega})$, as sketched in Figure P6.43. Suppose the impulse response is modulated with the sequence $(-1)^n$ to obtain $h_{\mathrm{hp}}[n] = (-1)^n h_{\mathrm{lp}}[n]$.

(a) Determine and sketch $H_{\mathrm{hp}}(e^{j\omega})$ in terms of $H_{\mathrm{lp}}(e^{j\omega})$. Show in particular that, for $H_{\mathrm{lp}}(e^{j\omega})$ as shown in Figure P6.43, $H_{\mathrm{hp}}(e^{j\omega})$ corresponds to a highpass filter.

(b) Show that modulation of the impulse response of a discrete-time highpass filter by $(-1)^n$ will transform it to a lowpass filter.

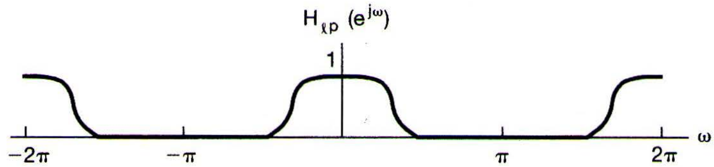


Figure P6.43

6.44. A discrete-time system is implemented as shown in Figure P6.44. The system S shown in the figure is an LTI system with impulse response $h_{\mathrm{lp}}[n]$.

(a) Show that the overall system is time invariant.

(b) If $h_{\mathrm{lp}}[n]$ is a lowpass filter, what type of filter does the system of the figure implement?

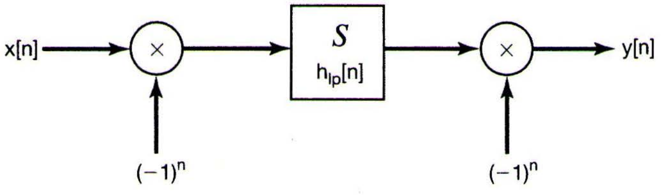


Figure P6.44

6.45. Consider the following three frequency responses for causal and stable third-order LTI systems. By utilizing the properties of first- and second-order systems discussed in Section 6.6, determine whether or not the impulse response of each of the third-order systems is oscillatory. (*Note:* You should be able to answer this question without taking the inverse Fourier transforms of the frequency responses of the third-order systems.)

$$H_1(e^{j\omega}) = \frac{1}{(1-\frac{1}{2}e^{-j\omega})(1-\frac{1}{3}e^{-j\omega})(1-\frac{1}{4}e^{-j\omega})},$$

$$H_2(e^{j\omega}) = \frac{1}{(1+\frac{1}{2}e^{-j\omega})(1-\frac{1}{3}e^{-j\omega})(1-\frac{1}{4}e^{-j\omega})},$$

$$H_3(e^{j\omega}) = \frac{1}{(1-\frac{1}{2}e^{-j\omega})(1-\frac{3}{4}e^{-j\omega}+\frac{9}{16}e^{-j2\omega})}.$$

6.46. Consider a causal, nonrecursive (FIR) filter whose real-valued impulse response $h[n]$ is zero for $n \geq N$.

(a) Assuming that N is odd, show that if $h[n]$ is symmetric about $(N-1)/2$ (i.e., if $h[(N-1)/2+n] = h[(N-1)/2-n]$), then

$$H(e^{j\omega}) = A(\omega)e^{-j[(N-1)/2]\omega},$$

where $A(\omega)$ is a real-valued function of ω. We conclude that the filter has linear phase.

(b) Give an example of the impulse response $h[n]$ of a causal, linear-phase FIR filter such that $h[n] = 0$ for $n \geq 5$ and $h[n] \neq 0$ for $0 \leq n \leq 4$.

(c) Assuming that N is even, show that if $h[n]$ is symmetric about $(N-1)/2$ (i.e., if $h[(N/2)+n] = h[N/2-n-1]$), then

$$H(e^{j\omega}) = A(\omega)e^{-j[(N-1)/2]\omega},$$

where $A(\omega)$ is a real-valued function of ω.

(d) Give an example of the impulse response $h[n]$ of a causal, linear-phase FIR filter such that $h[n] = 0$ for $n \geq 4$ and $h[n] \neq 0$ for $0 \leq n \leq 3$.

6.47. A three-point symmetric moving average, referred to as a weighted moving average, is of the form

$$y[n] = b\{ax[n-1] + x[n] + ax[n+1]\}. \qquad \text{(P6.47–1)}$$

(a) Determine, as a function of a and b, the frequency response $H(e^{j\omega})$ of the three-point moving average in eq. (P6.47–1).

(b) Determine the scaling factor b such that $H(e^{j\omega})$ has unity gain at zero frequency.

(c) In many time-series analysis problems, a common choice for the coefficient a in the weighted moving average in eq. (P6.47–1) is $a = 1/2$. Determine and sketch the frequency response of the resulting filter.

6.48. Consider a four-point, moving-average, discrete-time filter for which the difference equation is

$$y[n] = b_0x[n] + b_1x[n-1] + b_2x[n-2] + b_3x[n-2].$$

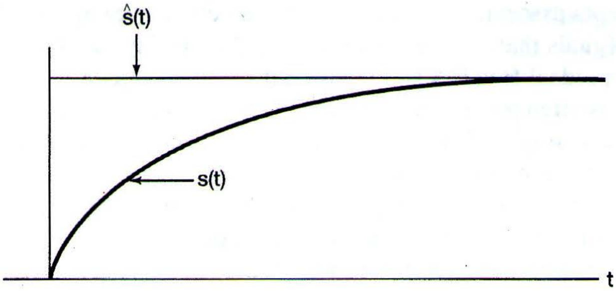


Figure P6.49

involving the analysis of complex systems having a few dominant time constants and other very small time constants. In order to reduce the complexity of the model of the system to be analyzed, one often can simplify the fast parts of the system. That is, suppose we regard a complex system as a parallel interconnection of first- and second-order systems. Suppose also that one of these subsystems, with impulse response $h(t)$ and step response $s(t)$, is fast—that is, that $s(t)$ settles to its final value $s(\infty)$ very quickly. Then we can approximate this subsystem by the subsystem that settles to the same final value *instantaneously*. That is, if $\hat{s}(t)$ is the step response to our approximation, then

$$\hat{s}(t) = s(\infty)u(t).$$

This is illustrated in Figure P6.49. Note that the impulse response of the approximate system is then

$$\hat{h}(t) = s(\infty)\delta(t),$$

which indicates that the approximate system is *memoryless*.

Consider again the causal LTI system described by eq. (P6.49–1) and, in particular, the representation of it as a parallel interconnection of two first-order systems, as described in part (b). Use the method just outlined to replace the faster of the two subsystems by a memoryless system. What is the differential equation that then describes the resulting overall system? What is the frequency response of this system? Sketch $|H(j\omega)|$ (not log $|H(j\omega)|$) and $\sphericalangle H(j\omega)$ for both the original and approximate systems. Over what range of frequencies are these frequency responses nearly equal? Sketch the step responses for both systems. Over what range of time are the step responses nearly equal? From your plots, you will see some of the similarities *and* differences between the original system and its approximation. The utility of an approximation such as this depends upon the specific application. In particular, one must take into account both how widely separated the different time constants are and also the nature of the inputs to be considered. As you will see from your answers in this part of the problem, the frequency response of the approximate system is essentially the same as the frequency response of the original system at low frequencies. That is, when the fast parts of the system are sufficiently fast compared to the rate of fluctuation of the input, the approximation becomes useful.

6.50. The concepts associated with frequency-selective filtering are often used to separate two signals that have been added together. If the spectra of the two signals do not overlap, ideal frequency-selective filters are desirable. However, if the spectra overlap, it is often preferable to design the filter to have a gradual transition between passband and stopband. In this problem, we explore one approach for determining the frequency response of a filter to be used for separating signals with overlapping spectra. Let $x(t)$ denote a composite continuous-time signal consisting of the sum of two signals $s(t) + w(t)$. As indicated in Figure P6.50(a), we would like to design an LTI filter to recover $s(t)$ from $x(t)$. The filter's frequency response $H(j\omega)$ is to be chosen so that, in some sense, $y(t)$ is a "good" approximation to $s(t)$.

Let us define a measure of the error between $y(t)$ and $s(t)$ at each frequency ω as

$$\epsilon(\omega) \triangleq |S(j\omega) - Y(j\omega)|^2,$$

where $S(j\omega)$ and $Y(j\omega)$ are the Fourier transforms of $s(t)$ and $y(t)$, respectively.

(a) Express $\epsilon(\omega)$ in terms of $S(j\omega)$, $H(j\omega)$, and $W(j\omega)$, where $W(j\omega)$ is the Fourier transform of $w(t)$.

(b) Let us restrict $H(j\omega)$ to be real, so that $H(j\omega) = H^*(j\omega)$. By setting the derivative of $\epsilon(\omega)$ with respect to $H(j\omega)$ to be zero, determine the $H(j\omega)$ required to minimize the error $\epsilon(\omega)$.

(c) Show that if the spectra of $S(j\omega)$ and $W(j\omega)$ are non-overlapping, the result in part (b) reduces to an ideal frequency-selective filter.

(d) From your result in part (b), determine and sketch $H(j\omega)$ if $S(j\omega)$ and $W(j\omega)$ are as shown in Figure P6.50(b).

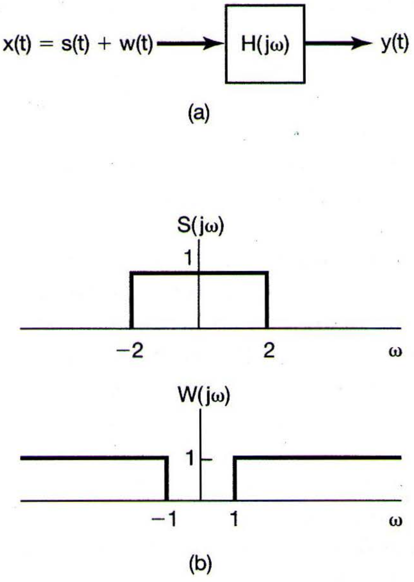


Figure P6.50

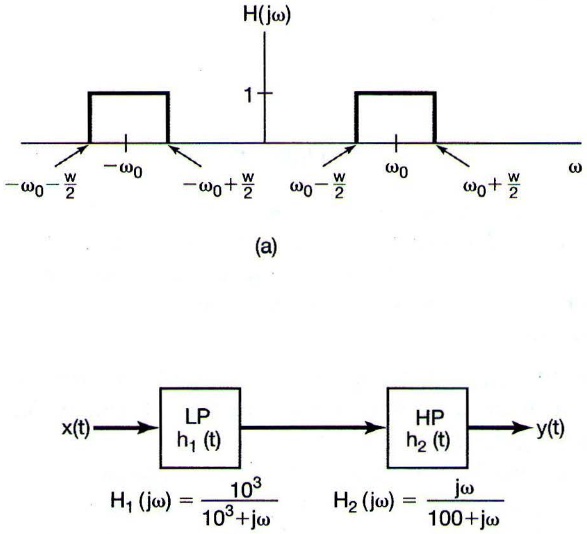
6.51. An ideal bandpass filter is a bandpass filter that passes only a range of frequencies, without any change in amplitude or phase. As shown in Figure P6.51(a), let the passband be

$$\omega_0 - \frac{W}{2} \leq |\omega| \leq \omega_0 + \frac{W}{2}.$$

(a) What is the impulse response $h(t)$ of this filter?

(b) We can approximate an ideal bandpass filter by cascading a first-order lowpass and a first-order highpass filter, as shown in Figure P6.51(b). Sketch the Bode diagrams for each of the two filters $H_1(j\omega)$ and $H_2(j\omega)$.

(c) Determine the Bode diagram for the overall bandpass filter in terms of your results from part (b).


H(jω)
1
$-\omega_0-\frac{W}{2}$
$-\omega_0$
$-\omega_0+\frac{W}{2}$
$\omega_0-\frac{W}{2}$
ω_0
$\omega_0+\frac{W}{2}$
ω
x(t)
LP
$h_1(t)$
HP
$h_2(t)$
y(t)
$H_1(j\omega) = \frac{10^3}{10^3+j\omega}$
$H_2(j\omega) = \frac{j\omega}{100+j\omega}$

(a)

(b)

Figure P6.51

6.52. In Figure P6.52(a), we show the magnitude of the frequency response for an ideal continuous-time differentiator. A nonideal differentiator would have a frequency response that is some approximation to the frequency response in the figure.

(a) Consider a nonideal differentiator with frequency response $G(j\omega)$ for which $|G(j\omega)|$ is constrained to be within $\pm 10\%$ of the magnitude of the frequency response of the ideal differentiator at all frequencies; that is,

$$-0.1|H(j\omega)| \leq [|G(j\omega)| - |H(j\omega)|] \leq 0.1|H(j\omega)|.$$

Sketch the region in a plot of $G(j\omega)$ vs. ω where $|G(j\omega)|$ must be confined to meet this specification.

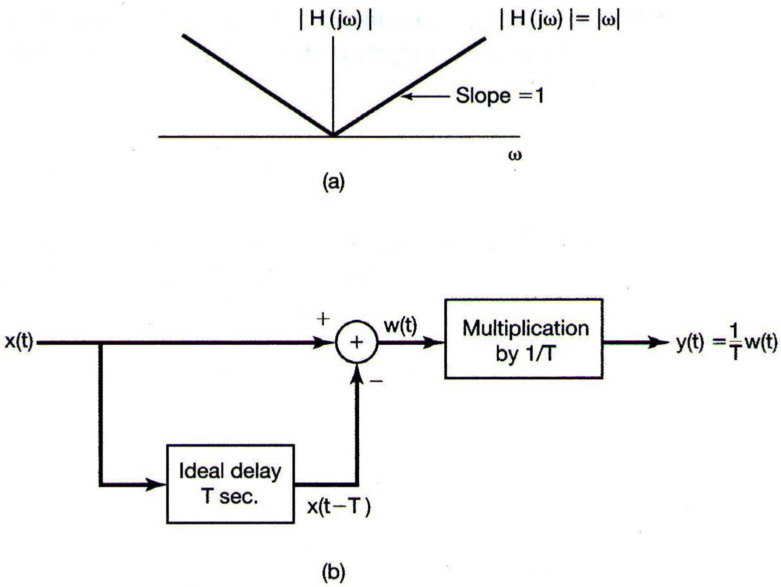


Figure P6.52

(b) The system in Figure P6.52(b), incorporating an ideal delay of T seconds, is sometimes used to approximate a continuous-time differentiator. For $T = 10^{-2}$ second, determine the frequency range over which the magnitude of the frequency response of the system in the figure is within ±10% of that for an ideal differentiator.

6.53. In many filtering applications, it is often undesirable for the step response of a filter to overshoot its final value. In processing pictures, for example, the overshoot in the step response of a linear filter may produce flare—that is, an increase in intensity—at sharp boundaries. It is possible, however, to eliminate overshoot by requiring that the impulse response of the filter be positive for all time.

Show that if $h(t)$, the impulse response of a continuous-time LTI filter, is always greater than or equal to zero, the step response of the filter is a monotonically nondecreasing function and therefore will not have overshoot.

6.54. By means of a specific filter design procedure, a nonideal continuous-time lowpass filter with frequency response $H_0(j\omega)$, impulse response $h_0(t)$, and step response $s_0(t)$ has been designed. The cutoff frequency of the filter is at $\omega = 2\pi \times 10^2$ rad/sec, and the step response rise time, defined as the time required for the step response to go from 10% of its final value to 90% of its final value, is $\tau_r = 10^{-2}$ second. From this design, we can obtain a new filter with an arbitrary cutoff frequency ω_c by the use of frequency scaling. The frequency response of the resulting filter is then of the form

$$H_{\text{lp}}(j\omega) = H_0(ja\omega),$$

where a is an appropriate scale factor.

(a) Determine the scale factor a such that $H_{\text{lp}}(j\omega)$ has a cutoff frequency of ω_c.

(b) Determine the impulse response $h_{\text{lp}}(t)$ of the new filter in terms of ω_c and $h_0(t)$.

(c) Determine the step response $s_{\text{lp}}(t)$ of the new filter in terms of ω_c and $s_0(t)$.

(d) Determine and sketch the rise time of the new filter as a function of its cutoff frequency ω_c.

This is one illustration of the trade-off between time-domain and frequency-domain characteristics. In particular, as the cutoff frequency decreases, the rise time tends to increase.

6.55. The square of the magnitude of the frequency response of a class of continuous-time lowpass filters, known as Butterworth filters, is

$$|B(j\omega)|^2 = \frac{1}{1+(\omega/\omega_c)^{2N}}.$$

Let us define the passband edge frequency ω_p as the frequency below which $|B(j\omega)|^2$ is greater than one-half of its value at $\omega = 0$; that is,

$$|B(j\omega)|^2 \geq \frac{1}{2}|B(j0)|^2, \quad |\omega| < \omega_p.$$

Now let us define the stopband edge frequency ω_s as the frequency above which $|B(j\omega)|^2$ is less than 10^{-2} of its value at $\omega = 0$; that is,

$$|B(j\omega)|^2 \leq 10^{-2}|B(j0)|^2, \quad |\omega| > \omega_s.$$

The transition band is then the frequency range between ω_p and ω_s. The ratio ω_s/ω_p is referred to as the transition ratio.

For fixed ω_p, and making reasonable approximations, determine and sketch the transition ratio as a function of N for the class of Butterworth filters.

6.56. In this problem, we explore some of the filtering issues involved in the commercial version of a typical system that is used in most modern cassette tape decks to reduce noise. The primary source of noise is the high-frequency hiss in the tape playback process, which, in some part, is due to the friction between the tape and the playback head. Let us assume that the noise hiss that is added to the signal upon playback has the spectrum of Figure P6.56(a) when measured in decibels, with 0 dB equal to the signal level at 100 Hz. The spectrum $S(j\omega)$ of the signal has the shape shown in Figure P6.56(b).

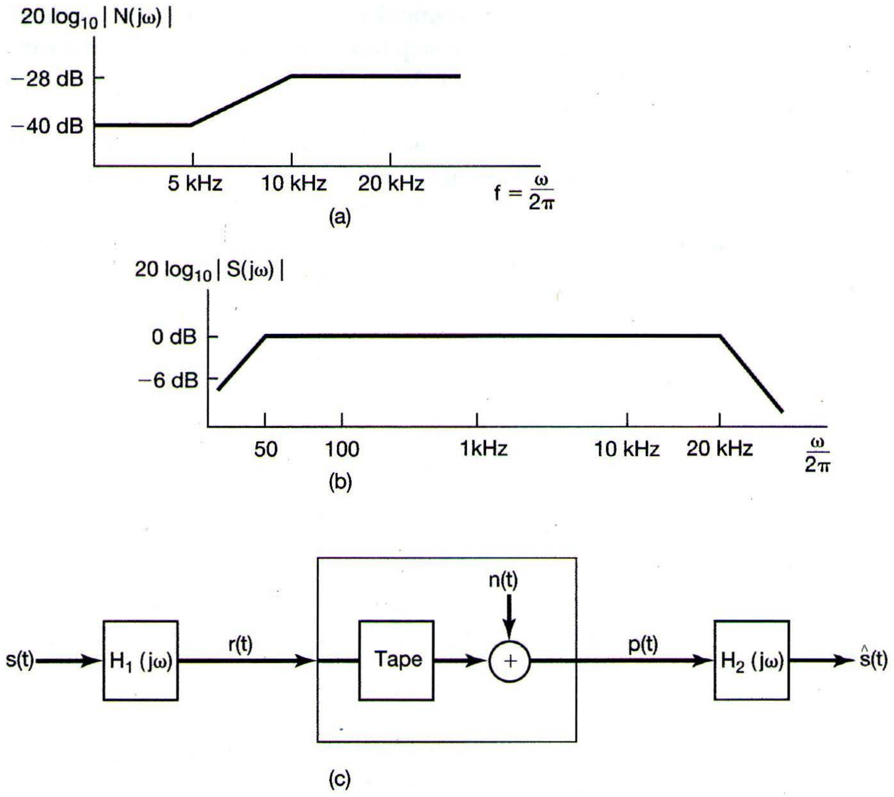
The system that we analyze has a filter $H_1(j\omega)$ which conditions the signal $s(t)$ before it is recorded. Upon playback, the hiss $n(t)$ is added to the signal. The system is represented schematically in Figure P6.56(c).

Suppose we would like our overall system to have a signal-to-noise ratio of 40 dB over the frequency range $50 \text{ Hz} < \omega/2\pi < 20 \text{ kHz}$.

(a) Determine the transfer characteristic of the filter $H_1(j\omega)$. Sketch the Bode plot of $H_1(j\omega)$.

(b) If we were to listen to the signal $p(t)$, assuming that the playback process does nothing more than add hiss to the signal, how do you think it would sound?

(c) What should the Bode plot and transfer characteristic of the filter $H_2(j\omega)$ be in order for the signal $\hat{s}(t)$ to sound similar to $s(t)$?



(a)

(b)

(c)

Figure P6.56

6.57. Show that if $h[n]$, the impulse response of a discrete-time LTI filter, is always greater than or equal to zero, the step response of the filter is a monotonically nondecreasing function and therefore will not have overshoot.

6.58. In the design of either analog or digital filters, we often approximate a specified magnitude characteristic without particular regard to the phase. For example, standard design techniques for lowpass and bandpass filters are typically derived from a consideration of the magnitude characteristics only.

In many filtering problems, one would ideally like the phase characteristics to be zero or linear. For causal filters, it is impossible to have zero phase. However, for many digital filtering applications, it is not necessary that the unit sample response of the filter be zero for $n < 0$ if the processing is not to be carried out in real time.

One technique commonly used in digital filtering when the data to be filtered are of finite duration and stored, for example, on a disc or magnetic tape is to process the data forward and then backward through the same filter.

Let $h[n]$ be the unit sample response of a causal filter with an arbitrary phase characteristic. Assume that $h[n]$ is real, and denote its Fourier transform by $H(e^{j\omega})$. Let $x[n]$ be the data that we want to filter. The filtering operation is performed as follows:

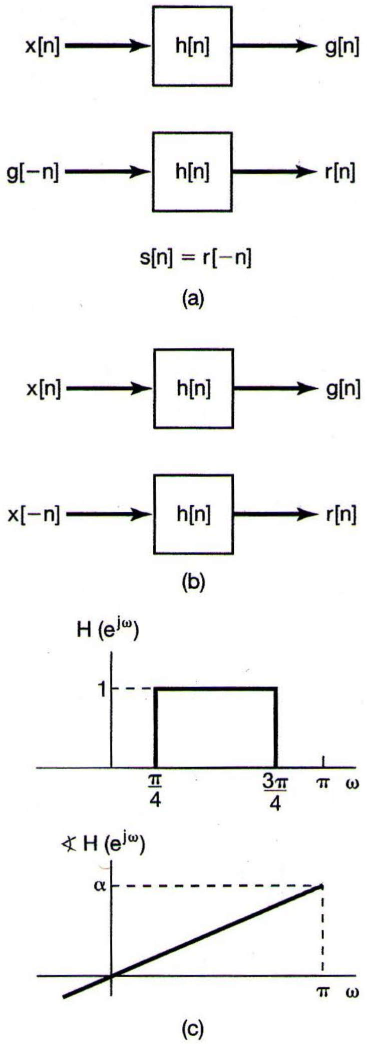


Figure P6.58

(a) *Method A:* Process $x[n]$ to get $s[n]$, as indicated in Figure P6.58(a).

1. Determine the overall unit sample response $h_1[n]$ that relates $x[n]$ and $s[n]$, and show that it has zero phase characteristic.
2. Determine $|H_1(e^{j\omega})|$ and express it in terms of $|H(e^{j\omega})|$ and $\sphericalangle H(e^{j\omega})$.

(b) *Method B:* Process $x[n]$ through the filter $h[n]$ to get $g[n]$ [Figure P6.58(b)]. Also, process $x[n]$ backward through $h[n]$ to get $r[n]$. The output $y[n]$ is taken to be the sum of $g[n]$ and $r[-n]$. The composite set of operations can be represented by a filter with input $x[n]$, output $y[n]$, and unit sample response $h_2[n]$.

1. Show that the composite filter $h_2[n]$ has zero phase characteristic.
2. Determine $|H_2(e^{j\omega})|$, and express it in terms of $|H(e^{j\omega})|$ and $\sphericalangle H(e^{j\omega})$.

(c) Suppose that we are given a sequence of finite duration on which we would like to perform bandpass, zero-phase filtering. Furthermore, assume that we are given the bandpass filter $h[n]$ with frequency response as specified in Figure P6.58(c) and with magnitude characteristic that we desire, but with linear phase. To achieve zero phase, we could use either of the preceding methods, A or B. Determine and sketch $|H_1(e^{j\omega})|$ and $|H_2(e^{j\omega})|$. From these results, which method would you use to achieve the desired bandpass filtering operation? Explain why. More generally, if $h[n]$ has the desired magnitude, but a nonlinear phase characteristic, which method is preferable to achieve a zero phase characteristic?

6.59. Let $h_d[n]$ denote the unit sample response of a desired ideal system with frequency response $H_d(e^{j\omega})$, and let $h[n]$ denote the unit sample response for an FIR system of length N and with frequency response $H(e^{j\omega})$. In this problem, we show that a rectangular window of length N samples applied to $h_d[n]$ will produce a unit sample response $h[n]$ such that the mean square error

$$\epsilon^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |H_d(e^{j\omega}) - H(e^{j\omega})|^2\, d\omega$$

is minimized.

(a) The error function $E(e^{j\omega}) = H_d(e^{j\omega}) - H(e^{j\omega})$ can be expressed as the power series

$$E(e^{j\omega}) = \sum_{n=-\infty}^{\infty} e[n]e^{-j\omega n}.$$

Find the coefficients $e[n]$ in terms of $h_d[n]$ and $h[n]$.

(b) Using Parseval's relation, express the mean square error ϵ^2 in terms of the coefficients $e[n]$.

(c) Show that for a unit sample response $h[n]$ of length N samples, ϵ^2 is minimized when

$$h[n] = \begin{cases} h_d[n], & 0 \le n \le N-1 \\ 0, & \text{otherwise} \end{cases}.$$

That is, simple truncation gives the best mean square approximation to a desired frequency response for a fixed value of N.

6.60. In Problem 6.50, we considered one specific criterion for determining the frequency response of a continuous-time filter that would recover a signal from the sum of two signals when their spectra overlapped in frequency. For the discrete-time case, develop the result corresponding to that obtained in part (b) of Problem 6.50.

6.61. In many situations we have available an analog or digital filter module, such as a basic hardware element or computer subroutine. By using the module repetitively or by combining identical modules, it is possible to implement a new filter with improved passband or stopband characteristics. In this and the next problem, we consider two procedures for doing just that. Although the discussion is phrased in terms of discrete-time filters, much of it applies directly to continuous-time filters as well.

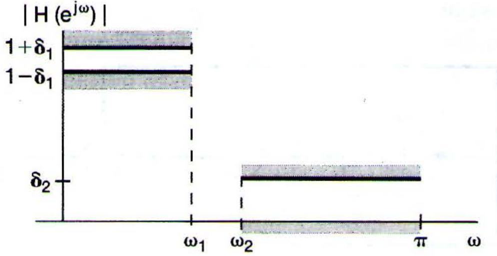


Figure P6.61

Consider a lowpass filter with frequency response $H(e^{j\omega})$ for which $|H(e^{j\omega})|$ falls within the tolerance limits shown in Figure P6.61; that is,

$$1-\delta_1 \le |H(e^{j\omega})| \le 1+\delta_1, \quad 0 \le \omega \le \omega_1,$$
$$0 \le |H(e^{j\omega})| \le \delta_2, \quad \omega_2 \le \omega \le \pi.$$

A new filter with frequency response $G(e^{j\omega})$ is formed by cascading two identical filters, both with frequency response $H(e^{j\omega})$.

(a) Determine the tolerance limits on $|G(e^{j\omega})|$.

(b) Assuming that $H(e^{j\omega})$ is a good approximation to a lowpass filter, so that $\delta_1 << 1$ and $\delta_2 << 1$, determine whether the passband ripple for $G(e^{j\omega})$ is larger or smaller than the passband ripple for $H(e^{j\omega})$. Also, determine whether the stopband ripple for $G(e^{j\omega})$ is larger or smaller than the stopband ripple for $H(e^{j\omega})$.

(c) If N identical filters with frequency response $H(e^{j\omega})$ are cascaded to obtain a new frequency response $G(e^{j\omega})$, then, again assuming that $\delta_1 << 1$ and $\delta_2 << 1$, determine the approximate tolerance limits on $|G(e^{j\omega})|$.

6.62. In Problem 6.61, we considered one method for using a basic filter module repetitively to implement a new filter with improved characteristics. Let us now consider an alternative approach, proposed by J. W. Tukey in the book, *Exploratory Data Analysis* (Reading, MA: Addison-Wesley Publishing Co., Inc., 1976). The procedure is shown in block diagram form in Figure P6.62(a).

(a) Suppose that $H(e^{j\omega})$ is real and has a passband ripple of $\pm\delta_1$ and a stopband ripple of $\pm\delta_2$ (i.e., $H(e^{j\omega})$ falls within the tolerance limits indicated in Figure P6.62(b)). The frequency response $G(e^{j\omega})$ of the overall system in Figure P6.62(a) falls within the tolerance limits indicated in Figure P6.62(c). Determine A, B, C, and D in terms of δ_1 and δ_2.

(b) If $\delta_1 << 1$ and $\delta_2 << 1$, what is the approximate passband ripple and stopband ripple associated with $G(e^{j\omega})$? Indicate in particular whether the passband ripple for $G(e^{j\omega})$ is larger or smaller than the passband ripple for $H(e^{j\omega})$. Also, indicate whether the stopband ripple for $G(e^{j\omega})$ is larger or smaller than the stopband ripple for $H(e^{j\omega})$.

(c) In parts (a) and (b), we assumed that $H(e^{j\omega})$ is real. Now consider $H(e^{j\omega})$ to have the more general form

$$H(e^{j\omega}) = H_1(e^{j\omega})e^{j\theta(\omega)},$$

where $H_1(e^{j\omega})$ is real and $\theta(\omega)$ is an unspecified phase characteristic. If $|H(e^{j\omega})|$

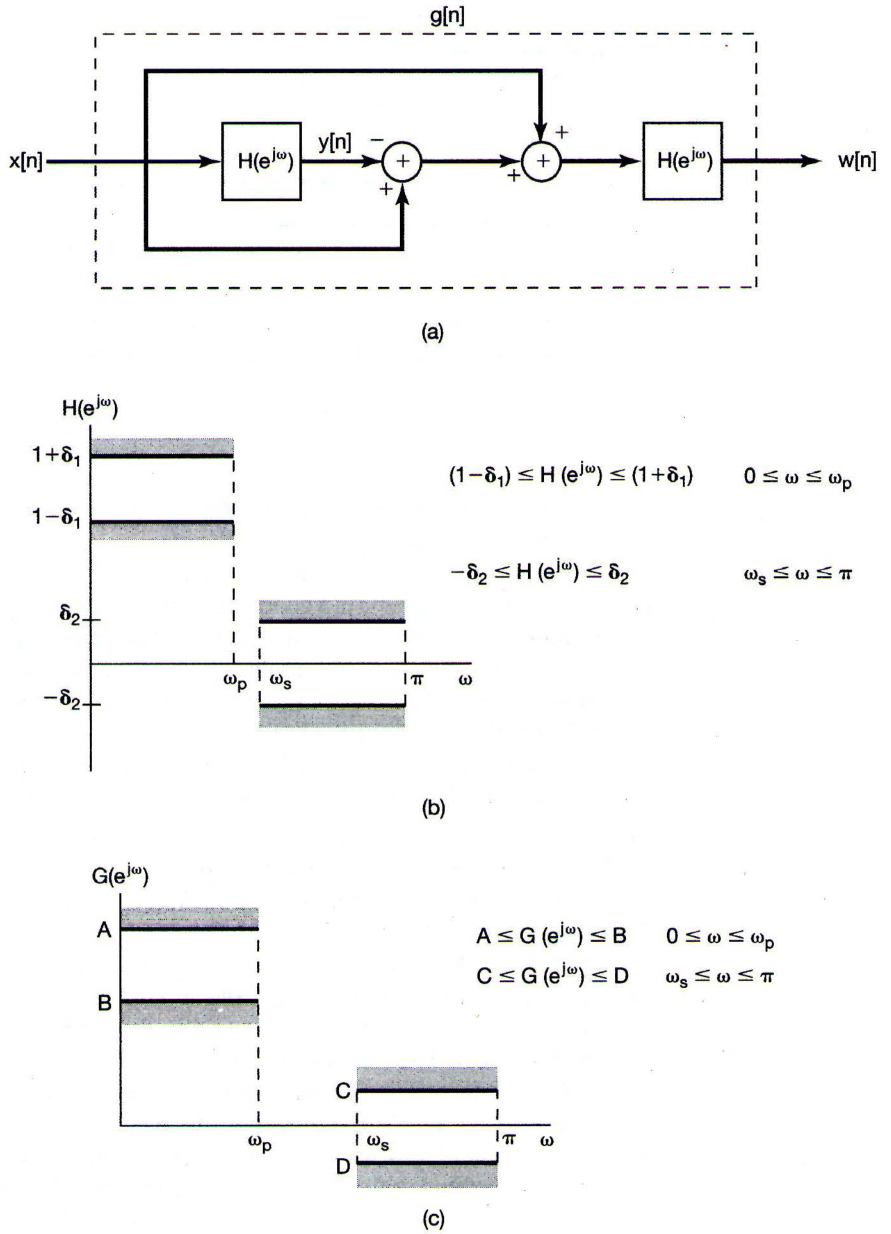

g[n]
x[n]
H(e^{jω})
y[n]
−
+
+
+
+
H(e^{jω})
w[n]
(a)
H(e^{jω})
1+δ₁
1−δ₁
δ₂
−δ₂
ωp
ωs
π
ω
(1−δ₁) ≤ H (e^{jω}) ≤ (1+δ₁) 0 ≤ ω ≤ ωp
−δ₂ ≤ H (e^{jω}) ≤ δ₂ ωs ≤ ω ≤ π
(b)
G(e^{jω})
A
B
C
D
ωp
ωs
π
ω
A ≤ G (e^{jω}) ≤ B 0 ≤ ω ≤ ωp
C ≤ G (e^{jω}) ≤ D ωs ≤ ω ≤ π
(c)

Figure P6.62

is a reasonable approximation to an ideal lowpass filter, will $|G(e^{j\omega})|$ necessarily be a reasonable approximation to an ideal lowpass filter?

(d) Now assume that $H(e^{j\omega})$ is an FIR linear-phase lowpass filter, so that

$$H(e^{j\omega}) = H_1(e^{j\omega})e^{jM\omega},$$

where $H_1(e^{j\omega})$ is real and M is an integer. Show how to modify the system in Figure P6.62(a) so that the overall system will approximate a lowpass filter.

6.63. In the design of digital filters, we often choose a filter with a specified magnitude characteristic that has the shortest duration. That is, the impulse response, which is the inverse Fourier transform of the complex frequency spectrum, should be as narrow as possible. Assuming that $h[n]$ is real, we wish to show that if the phase $\theta(\omega)$ associated with the frequency response $H(e^{j\omega})$ is zero, the duration of the impulse response is minimal. Let the frequency response be expressed as

$$H(e^{j\omega}) = |H(e^{j\omega})|e^{j\theta(\omega)},$$

and let us consider the quantity

$$D = \sum_{n=-\infty}^{\infty} n^2h^2[n] = \sum_{n=-\infty}^{\infty} (nh[n])^2$$

to be a measure of the duration of the associated impulse response $h[n]$.

(a) Using the derivative property of the Fourier transform and Parseval's relation, express D in terms of $H(e^{j\omega})$.

(b) By expressing $H(e^{j\omega})$ in terms of its magnitude $|H(e^{j\omega})|$ and phase $\theta(\omega)$, use your result from part (a) to show that D is minimized when $\theta(\omega) = 0$.

6.64. For a discrete-time filter to be *causal* and to have exactly linear phase, its impulse response must be of finite length and consequently the difference equation must be nonrecursive. To focus on the insight behind this statement, we consider a particular case, that of a linear phase characteristic for which the slope of the phase is an integer. Thus, the frequency response is assumed to be of the form

$$H(e^{j\omega}) = H_r(e^{j\omega})e^{-jM\omega}, \; -\pi < \omega < \pi \qquad \text{(P6.64–1)}$$

where $H_r(e^{j\omega})$ is real and even.

Let $h[n]$ denote the impulse response of the filter with frequency response $H(e^{j\omega})$ and let $h_r[n]$ denote the impulse response of the filter with frequency response $H_r(e^{j\omega})$.

(a) By using the appropriate properties in Table 5.1, show that:

1. $h_r[n] = h_r[-n]$ (i.e., $h_r[n]$ is symmetric about $n = 0$).

2. $h[n] = h_r[n - M]$.

(b) Using your result in part (a), show that with $H(e^{j\omega})$ of the form shown in eq. (P6.64–1), $h[n]$ is symmetric about $n = M$, that is,

$$h[M + n] = h[M - n]. \qquad \text{(P6.64–2)}$$

(c) According to the result in part (b), the linear phase characteristic in eq. (P6.64–1) imposes a symmetry in the impulse response. Show that if $h[n]$ is causal and has the symmetry in eq. (P6.64–2), then

$$h[n] = 0, \qquad n < 0 \text{ and } n > 2M$$

(i.e., it must be of finite length).

6.65. For a class of discrete-time lowpass filters, known as Butterworth filters, the squared magnitude of the frequency response is given by

$$|B(e^{j\omega})|^2 = \frac{1}{1 + \left(\frac{\tan(\omega/2)}{\tan(\omega_c/2)}\right)^{2N}},$$

where ω_c is the cutoff frequency (which we shall take to be $\pi/2$) and N is the order of the filter (which we shall consider to be $N = 1$). Thus, we have

$$|B(e^{j\omega})|^2 = \frac{1}{1 + \tan^2(\omega/2)}.$$

(a) Using trigonometric identities, show that $|B(e^{j\omega})|^2 = \cos^2(\omega/2)$.

(b) Let $B(e^{j\omega}) = a\cos(\omega/2)$. For what complex values of a is $|B(e^{j\omega})|^2$ the same as in part (a)?

(c) Show that $B(e^{j\omega})$ from part (b) is the transfer function corresponding to a difference equation of the form

$$y[n] = \alpha x[n] + \beta x[n - \gamma].$$

Determine α, β, and γ.

6.66. In Figure P6.66(a) we show a discrete-time system consisting of a parallel combination of N LTI filters with impulse response $h_k[n]$, $k = 0, 1, \cdots, N - 1$. For any k, $h_k[n]$ is related to $h_0[n]$ by the expression

$$h_k[n] = e^{j(2\pi nk/N)} h_0[n].$$

(a) If $h_0[n]$ is an ideal discrete-time lowpass filter with frequency response $H_0(e^{j\omega})$ as shown in Figure P6.66(b), sketch the Fourier transforms of $h_1[n]$ and $h_{N-1}[n]$ for ω in the range $-\pi < \omega \leq +\pi$.

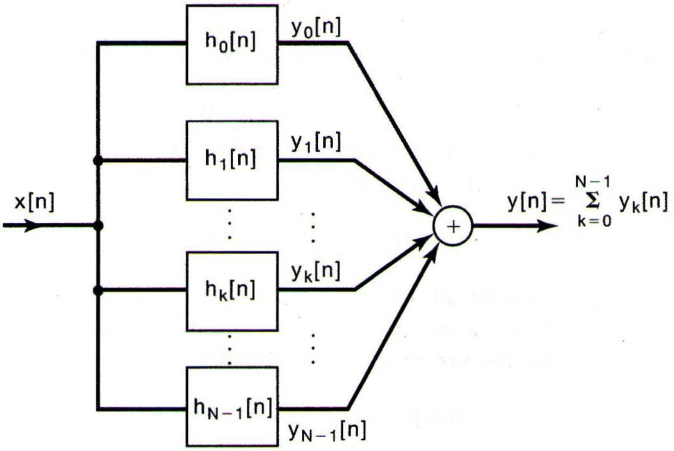


Figure P6.66a

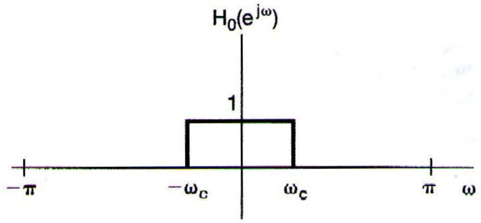


Figure P6.66b

(b) Determine the value of the cutoff frequency ω_c in Figure P6.66(b) in terms of N $(0 < \omega_c \leq \pi)$ such that the system of Figure P6.66(a) is an identity system; that is, $y[n] = x[n]$ for all n and any input $x[n]$.

(c) Suppose that $h[n]$ is no longer restricted to be an ideal lowpass filter. If $h[n]$ denotes the impulse response of the entire system in Figure P6.66(a) with input $x[n]$ and output $y[n]$, then $h[n]$ can be expressed in the form

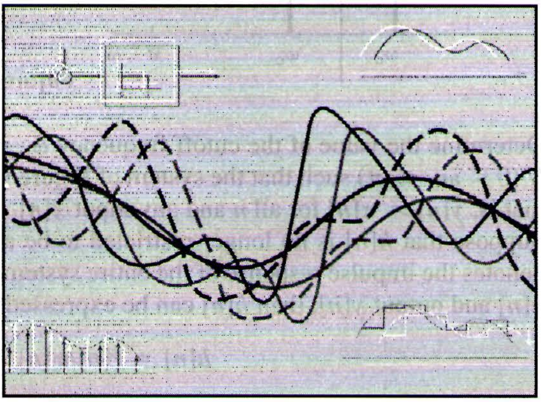
$$h[n] = r[n]h_0[n].$$

Determine and sketch $r[n]$.

(d) From your result of part (c), determine a necessary and sufficient condition on $h_0[n]$ to ensure that the overall system will be an identity system (i.e., such that for any input $x[n]$, the output $y[n]$ will be identical to $x[n]$). Your answer should not contain any sums.

7

SAMPLING



7.0 INTRODUCTION

Under certain conditions, a continuous-time signal can be completely represented by and recoverable from knowledge of its values, or *samples,* at points equally spaced in time. This somewhat surprising property follows from a basic result that is referred to as the *sampling theorem.* This theorem is extremely important and useful. It is exploited, for example, in moving pictures, which consist of a sequence of individual frames, each of which represents an instantaneous view (i.e., a sample in time) of a continuously changing scene. When these samples are viewed in sequence at a sufficiently fast rate, we perceive an accurate representation of the original continuously moving scene. As another example, printed pictures typically consist of a very fine grid of points, each corresponding to a sample of the spatially continuous scene represented in the picture. If the samples are sufficiently close together, the picture appears to be spatially continuous, although under a magnifying glass its representation in terms of samples becomes evident.

Much of the importance of the sampling theorem also lies in its role as a bridge between continuous-time signals and discrete-time signals. As we will see in this chapter, the fact that under certain conditions a continuous-time signal can be completely recovered from a sequence of its samples provides a mechanism for representing a continuous-time signal by a discrete-time signal. In many contexts, processing discrete-time signals is more flexible and is often preferable to processing continuous-time signals. This is due in large part to the dramatic development of digital technology over the past few decades, resulting in the availability of inexpensive, lightweight, programmable, and easily reproducible discrete-time systems. The concept of sampling, then, suggests an extremely attractive and widely employed method for using discrete-time system technology to implement continuous-time systems and process continuous-time signals: We exploit sampling to convert a continuous-time signal to a discrete-time signal, process the discrete-time signal using a discrete-time system, and then convert back to continuous time.

In the following discussion, we introduce and develop the concept of sampling and the process of reconstructing a continuous-time signal from its samples. In this discussion, we both identify the conditions under which a continuous-time signal can be exactly reconstructed from its samples and examine the consequences when these conditions are not satisfied. Following this, we explore the processing of continuous-time signals that have been converted to discrete-time signals through sampling. Finally, we examine the sampling of discrete-time signals and the related concepts of decimation and interpolation.

7.1 REPRESENTATION OF A CONTINUOUS-TIME SIGNAL BY ITS SAMPLES: THE SAMPLING THEOREM

In general, in the absence of any additional conditions or information, we would not expect that a signal could be uniquely specified by a sequence of equally spaced samples. For example, in Figure 7.1 we illustrate three different continuous-time signals, all of which have identical values at integer multiples of T; that is,

$$x_1(kT) = x_2(kT) = x_3(kT).$$

Clearly, an infinite number of signals can generate a given set of samples. As we will see, however, if a signal is band limited—i.e., if its Fourier transform is zero outside a finite band of frequencies—and if the samples are taken sufficiently close together in relation to the highest frequency present in the signal, then the samples *uniquely* specify the signal, and we can reconstruct it perfectly. This result, known as the *sampling theorem*, is of profound importance in the practical application of the methods of signal and system analysis.

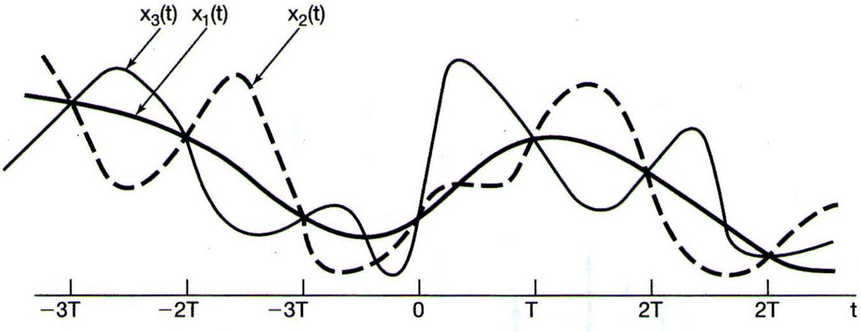


Figure 7.1 Three continuous-time signals with identical values at integer multiples of T.

7.1.1 Impulse-Train Sampling

In order to develop the sampling theorem, we need a convenient way in which to represent the sampling of a continuous-time signal at regular intervals. A useful way to do this is through the use of a periodic impulse train multiplied by the continuous-time signal $x(t)$ that we wish to sample. This mechanism, known as *impulse-train sampling,* is depicted in Figure 7.2. The periodic impulse train $p(t)$ is referred to as the *sampling function,* the period T as the *sampling period,* and the fundamental frequency of $p(t)$, $\omega_s = 2\pi/T$, as the *sampling frequency.* In the time domain,

$$x_p(t) = x(t)p(t), \tag{7.1}$$

where

$$p(t) = \sum_{n=-\infty}^{+\infty} \delta(t - nT). \tag{7.2}$$

Because of the sampling property of the unit impulse discussed in Section 1.4.2, we know that multiplying $x(t)$ by a unit impulse samples the value of the signal at the point at which the impulse is located; i.e., $x(t)\delta(t-t_0) = x(t_0)\delta(t-t_0)$. Applying this to eq. (7.1), we see, as illustrated in Figure 7.2, that $x_p(t)$ is an impulse train with the amplitudes of

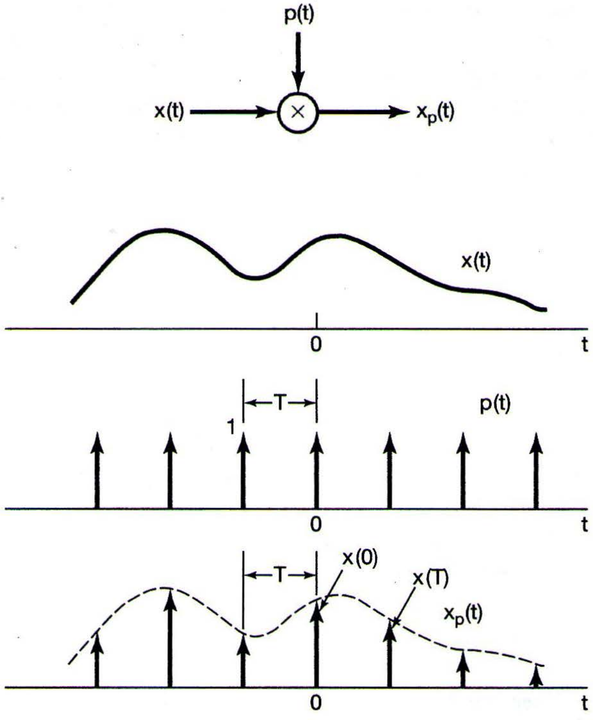


Figure 7.2 Impulse-train sampling.

the impulses equal to the samples of $x(t)$ at intervals spaced by T; that is,

$$x_p(t) = \sum_{n=-\infty}^{+\infty} x(nT)\delta(t - nT). \tag{7.3}$$

From the multiplication property (Section 4.5), we know that

$$X_p(j\omega) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} X(j\theta)P(j(\omega - \theta))d\theta. \tag{7.4}$$

and from Example 4.8,

$$P(j\omega) = \frac{2\pi}{T}\sum_{k=-\infty}^{+\infty} \delta(\omega - k\omega_s). \tag{7.5}$$

Since convolution with an impulse simply shifts a signal [i.e., $X(j\omega) * \delta(\omega - \omega_0) = X(j(\omega - \omega_0))$], it follows that

$$X_p(j\omega) = \frac{1}{T}\sum_{k=-\infty}^{+\infty} X(j(\omega - k\omega_s)). \tag{7.6}$$

That is, $X_p(j\omega)$ is a periodic function of ω consisting of a superposition of shifted replicas of $X(j\omega)$, scaled by $1/T$, as illustrated in Figure 7.3. In Figure 7.3(c), $\omega_M < (\omega_s - \omega_M)$, or equivalently, $\omega_s > 2\omega_M$, and thus there is no overlap between the shifted replicas of $X(j\omega)$, whereas in Figure 7.3(d), with $\omega_s < 2\omega_M$, there is overlap. For the case illustrated in Figure 7.3(c), $X(j\omega)$ is faithfully reproduced at integer multiples of the sampling frequency. Consequently, if $\omega_s > 2\omega_M$, $x(t)$ can be recovered exactly from $x_p(t)$ by means of

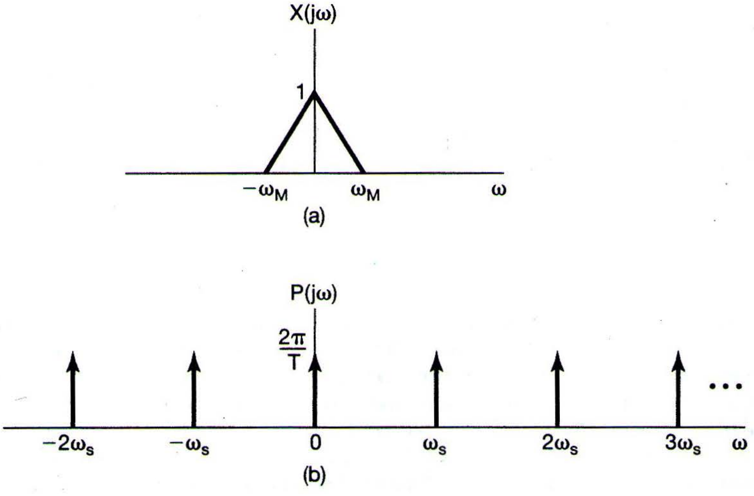


Figure 7.3 Effect in the frequency domain of sampling in the time domain: (a) spectrum of original signal; (b) spectrum of sampling function;

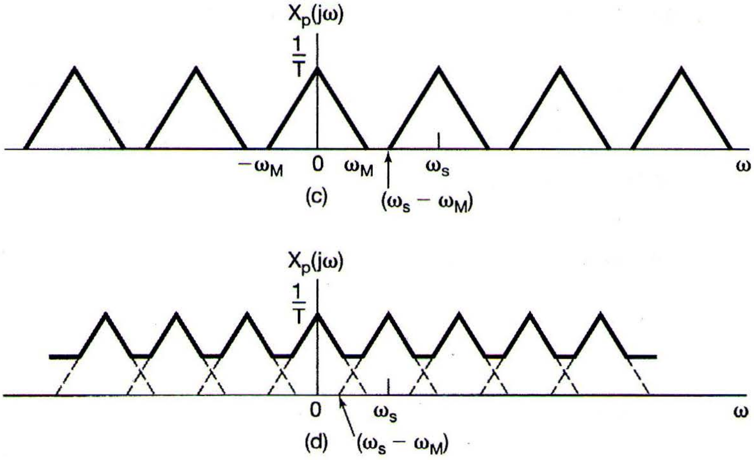


Figure 7.3 *Continued* (c) spectrum of sampled signal with $\omega_s > 2\omega_M$; (d) spectrum of sampled signal with $\omega_s < 2\omega_M$.

a lowpass filter with gain T and a cutoff frequency greater than ω_M and less than $\omega_s - \omega_M$, as indicated in Figure 7.4. This basic result, referred to as the *sampling theorem,* can be stated as follows: [1]

Sampling Theorem:
Let $x(t)$ be a band-limited signal with $X(j\omega) = 0$ for $|\omega| > \omega_M$. Then $x(t)$ is uniquely determined by its samples $x(nT)$, $n = 0, \pm 1, \pm 2, \ldots$, if

$$\omega_s > 2\omega_M,$$

where

$$\omega_s = \frac{2\pi}{T}.$$

Given these samples, we can reconstruct $x(t)$ by generating a periodic impulse train in which successive impulses have amplitudes that are successive sample values. This impulse train is then processed through an ideal lowpass filter with gain T and cutoff frequency greater than ω_M and less than $\omega_s - \omega_M$. The resulting output signal will exactly equal $x(t)$.

---

[1] The important and elegant sampling theorem was available for many years in a variety of forms in the mathematics literature. See, for example, J. M. Whittaker, "Interpolatory Function Theory," (New York: Stecher-Hafner Service Agency, 1964), chap. 4. It did not appear explicitly in the literature of communication theory until the publication in 1949 of the classic paper by Shannon entitled "Communication in the Presence of Noise" (*Proceedings of the IRE,* January 1949, pp. 10–21). However, H. Nyquist in 1928 and D. Gabor in 1946 had pointed out, based on the use of the Fourier Series, that $2TW$ numbers are sufficient to represent a function of duration T and highest frequency W. [H. Nyquist, "Certain Topics in Telegraph Transmission Theory," *AIEE Transactions,* 1928, p. 617; D. Gabor, "Theory of Communication,"*Journal of IEE* 93, no. 26 (1946), p. 429.]

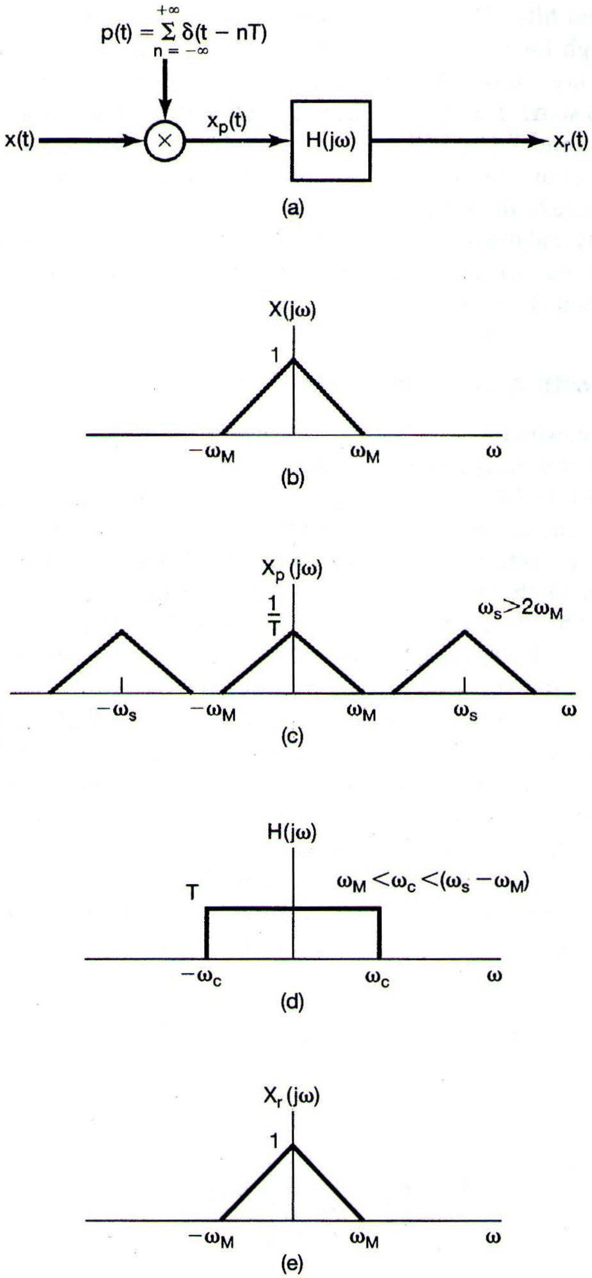

$p(t) = \sum_{n=-\infty}^{+\infty} \delta(t - nT)$
$x(t)$
$x_p(t)$
$H(j\omega)$
$x_r(t)$
(a)
$X(j\omega)$
1
$-\omega_M$
ω_M
ω
(b)
$X_p(j\omega)$
$\frac{1}{T}$
$\omega_s > 2\omega_M$
$-\omega_s$
$-\omega_M$
ω_M
ω_s
ω
(c)
$H(j\omega)$
T
$\omega_M < \omega_c < (\omega_s - \omega_M)$
$-\omega_c$
ω_c
ω
(d)
$X_r(j\omega)$
1
$-\omega_M$
ω_M
ω
(e)

Figure 7.4 Exact recovery of a continuous-time signal from its samples using an ideal lowpass filter: (a) system for sampling and reconstruction; (b) representative spectrum for $x(t)$; (c) corresponding spectrum for $x_p(t)$; (d) ideal lowpass filter to recover $X(j\omega)$ from $X_p(j\omega)$; (e) spectrum of $x_r(t)$.

The frequency $2\omega_M$, which, under the sampling theorem, must be exceeded by the sampling frequency, is commonly referred to as the *Nyquist rate.*[2]

As discussed in Chapter 6, ideal filters are generally not used in practice for a variety of reasons. In any practical application, the ideal lowpass filter in Figure 7.4 would be replaced by a nonideal filter $H(j\omega)$ that approximated the desired frequency characteristic accurately enough for the problem of interest (i.e., $H(j\omega) \simeq 1$ for $|\omega| < \omega_M$, and $H(j\omega) \simeq 0$ for $|\omega| > \omega_S - \omega_M$). Obviously, any such approximation in the lowpass filtering stage will lead to some discrepancy between $x(t)$ and $x_r(t)$ in Figure 7.4 or, equivalently, between $X(j\omega)$ and $X_r(j\omega)$. The particular choice of nonideal filter is then dictated by the acceptable level of distortion for the application under consideration. For convenience and to emphasize basic principles such as the sampling theorem, we will regularly assume the availability and use of ideal filters throughout this and the next chapter, with the understanding that in practice such a filter must be replaced by a nonideal filter designed to approximate the ideal characteristics accurately enough for the problem at hand.

---

[2] The frequency ω_M corresponding to one-half the Nyquist rate is often referred to as the *Nyquist frequency.*

7.1.2 Sampling with a Zero-Order Hold

The sampling theorem, which is most easily explained in terms of impulse-train sampling, establishes the fact that a band-limited signal is uniquely represented by its samples. In practice, however, narrow, large-amplitude pulses, which approximate impulses, are also relatively difficult to generate and transmit, and it is often more convenient to generate the sampled signal in a form referred to as a *zero-order hold.* Such a system samples $x(t)$ at a given instant and holds that value until the next instant at which a sample is taken, as illustrated in Figure 7.5. The reconstruction of $x(t)$ from the output of a zero-order hold can again be carried out by lowpass filtering. However, in this case, the required filter no longer has constant gain in the passband. To develop the required filter characteristic, we first note that the output $x_0(t)$ of the zero-order hold can in principle be generated by impulse-train sampling followed by an LTI system with a rectangular impulse response, as depicted in Figure 7.6. To reconstruct $x(t)$ from $x_0(t)$, we consider processing $x_0(t)$ with an LTI system with impulse response $h_r(t)$ and frequency response $H_r(j\omega)$. The cascade of this system with the system of Figure 7.6 is shown in Figure 7.7, where we wish to specify $H_r(j\omega)$ so that $r(t) = x(t)$. Comparing the system in Figure 7.7 with that in Figure 7.4, we see that $r(t) = x(t)$ if the cascade combination of $h_0(t)$ and $h_r(t)$ is the ideal lowpass filter $H(j\omega)$ used in Figure 7.4. Since, from Example 4.4 and the time-shifting property in Section 4.3.2,

$$H_0(j\omega) = e^{-j\omega T/2}\left[\frac{2\sin(\omega T/2)}{\omega}\right], \tag{7.7}$$

this requires that

$$H_r(j\omega) = \frac{e^{j\omega T/2}H(j\omega)}{\frac{2\sin(\omega T/2)}{\omega}}. \tag{7.8}$$

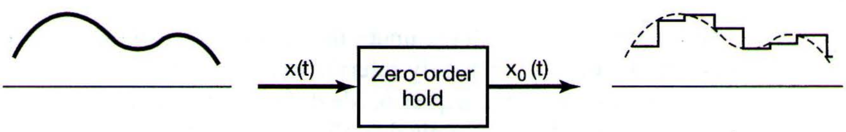
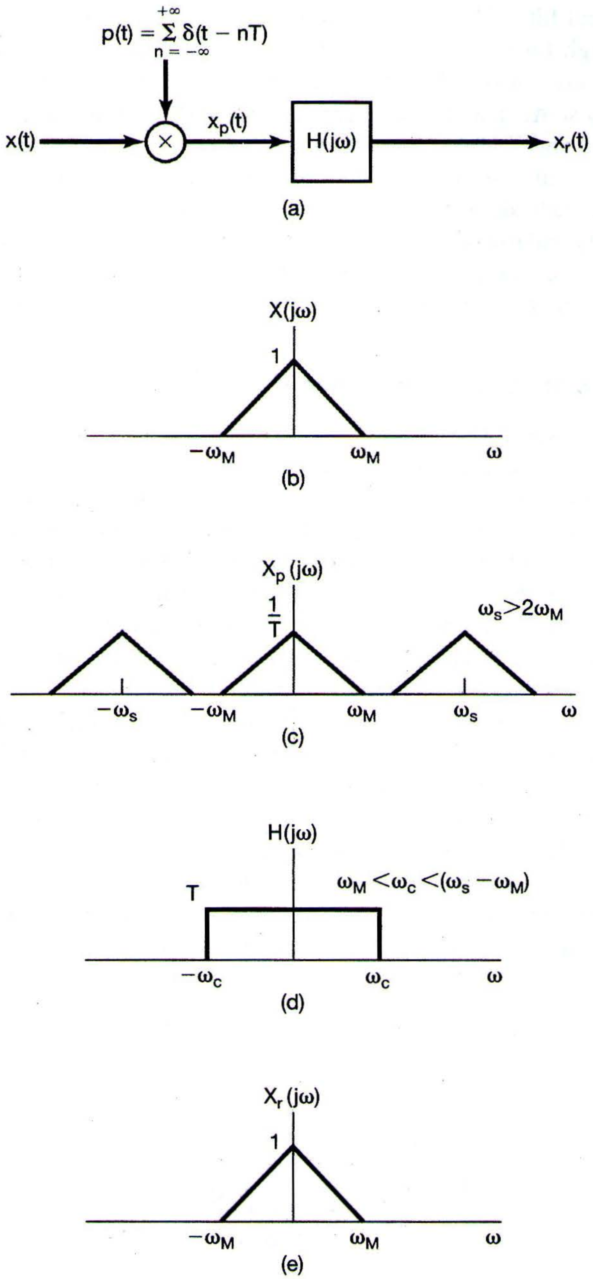


Figure 7.5 Sampling utilizing a zero-order hold.



$p(t) = \sum_{n=-\infty}^{+\infty} \delta(t - nT)$
x(t)
×
$x_p(t)$
$H(j\omega)$
$x_r(t)$
(a)
$X(j\omega)$
1
$-\omega_M$
ω_M
ω
(b)
$X_p(j\omega)$
$\frac{1}{T}$
$\omega_s > 2\omega_M$
$-\omega_s$
$-\omega_M$
ω_M
ω_s
ω
(c)
$H(j\omega)$
T
$\omega_M < \omega_c < (\omega_s - \omega_M)$
$-\omega_c$
ω_c
ω
(d)
$X_r(j\omega)$
1
$-\omega_M$
ω_M
ω
(e)

Figure 7.4 Exact recovery of a continuous-time signal from its samples using an ideal lowpass filter: (a) system for sampling and reconstruction; (b) representative spectrum for $x(t)$; (c) corresponding spectrum for $x_p(t)$; (d) ideal lowpass filter to recover $X(j\omega)$ from $X_p(j\omega)$; (e) spectrum of $x_r(t)$.

The frequency $2\omega_M$, which, under the sampling theorem, must be exceeded by the sampling frequency, is commonly referred to as the *Nyquist rate.*[2]

As discussed in Chapter 6, ideal filters are generally not used in practice for a variety of reasons. In any practical application, the ideal lowpass filter in Figure 7.4 would be replaced by a nonideal filter $H(j\omega)$ that approximated the desired frequency characteristic accurately enough for the problem of interest (i.e., $H(j\omega) \simeq 1$ for $|\omega| < \omega_M$, and $H(j\omega) \simeq 0$ for $|\omega| > \omega_S - \omega_M$). Obviously, any such approximation in the lowpass filtering stage will lead to some discrepancy between $x(t)$ and $x_r(t)$ in Figure 7.4 or, equivalently, between $X(j\omega)$ and $X_r(j\omega)$. The particular choice of nonideal filter is then dictated by the acceptable level of distortion for the application under consideration. For convenience and to emphasize basic principles such as the sampling theorem, we will regularly assume the availability and use of ideal filters throughout this and the next chapter, with the understanding that in practice such a filter must be replaced by a nonideal filter designed to approximate the ideal characteristics accurately enough for the problem at hand.

---

[2] The frequency ω_M corresponding to one-half the Nyquist rate is often referred to as the *Nyquist frequency.*

7.1.2 Sampling with a Zero-Order Hold

The sampling theorem, which is most easily explained in terms of impulse-train sampling, establishes the fact that a band-limited signal is uniquely represented by its samples. In practice, however, narrow, large-amplitude pulses, which approximate impulses, are also relatively difficult to generate and transmit, and it is often more convenient to generate the sampled signal in a form referred to as a *zero-order hold.* Such a system samples $x(t)$ at a given instant and holds that value until the next instant at which a sample is taken, as illustrated in Figure 7.5. The reconstruction of $x(t)$ from the output of a zero-order hold can again be carried out by lowpass filtering. However, in this case, the required filter no longer has constant gain in the passband. To develop the required filter characteristic, we first note that the output $x_0(t)$ of the zero-order hold can in principle be generated by impulse-train sampling followed by an LTI system with a rectangular impulse response, as depicted in Figure 7.6. To reconstruct $x(t)$ from $x_0(t)$, we consider processing $x_0(t)$ with an LTI system with impulse response $h_r(t)$ and frequency response $H_r(j\omega)$. The cascade of this system with the system of Figure 7.6 is shown in Figure 7.7, where we wish to specify $H_r(j\omega)$ so that $r(t) = x(t)$. Comparing the system in Figure 7.7 with that in Figure 7.4, we see that $r(t) = x(t)$ if the cascade combination of $h_0(t)$ and $h_r(t)$ is the ideal lowpass filter $H(j\omega)$ used in Figure 7.4. Since, from Example 4.4 and the time-shifting property in Section 4.3.2,

$$H_0(j\omega) = e^{-j\omega T/2}\left[\frac{2\sin(\omega T/2)}{\omega}\right], \tag{7.7}$$

this requires that

$$H_r(j\omega) = \frac{e^{j\omega T/2}H(j\omega)}{\frac{2\sin(\omega T/2)}{\omega}}. \tag{7.8}$$

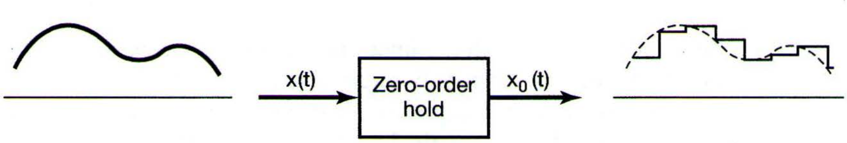


Figure 7.5 Sampling utilizing a zero-order hold.

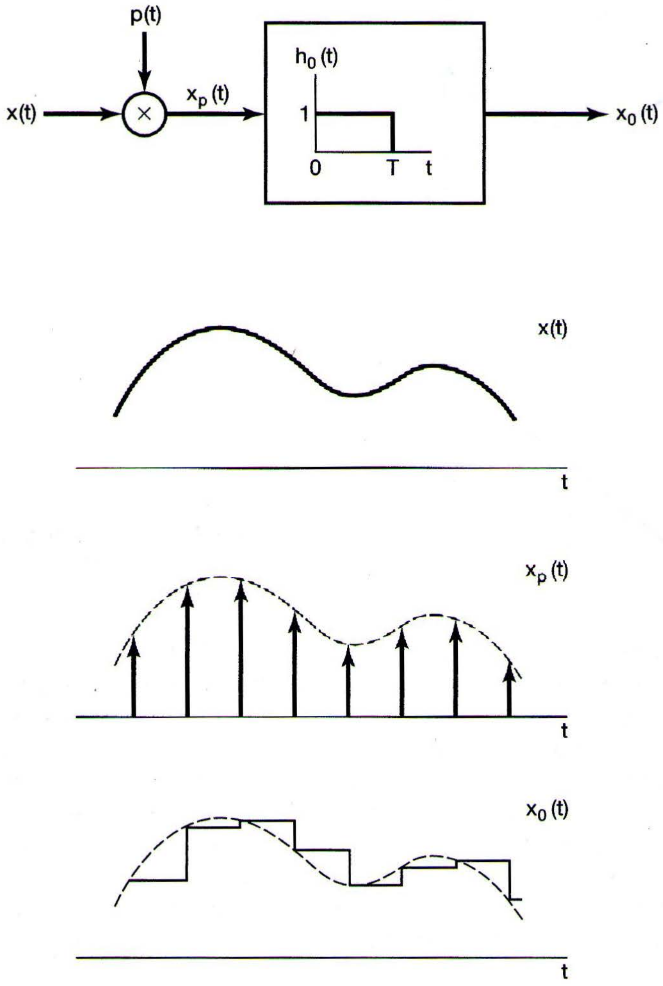

p(t)
x(t)
$x_p(t)$
$h_0(t)$
1
0
T
t
$x_0(t)$
x(t)
t
$x_p(t)$
t
$x_0(t)$
t

Figure 7.6 Zero-order hold as impulse-train sampling followed by an LTI system with a rectangular impulse response.

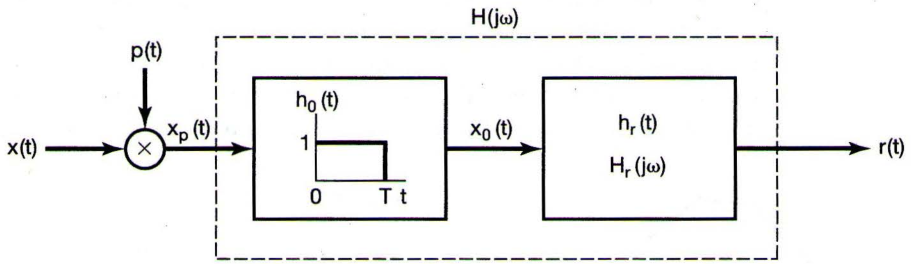

H(jω)
p(t)
x(t)
$x_p(t)$
$h_0(t)$
1
0
T t
$x_0(t)$
$h_r(t)$
$H_r(j\omega)$
r(t)

Figure 7.7 Cascade of the representation of a zero-order hold (Figure 7.6) with a reconstruction filter.

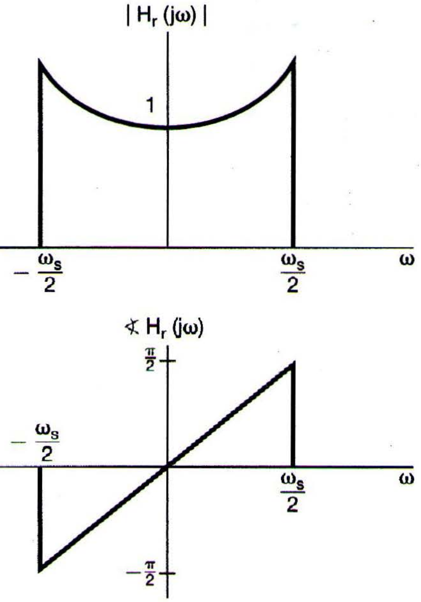


Figure 7.8 Magnitude and phase for the reconstruction filter for a zero-order hold.

For example, with the cutoff frequency of $H(j\omega)$ equal to $\omega_s/2$, the ideal magnitude and phase for the reconstruction filter following a zero-order hold is that shown in Figure 7.8.

Once again, in practice the frequency response in eq. (7.8) cannot be exactly realized, and thus an adequate approximation to it must be designed. In fact, in many situations, the output of the zero-order hold is considered an adequate approximation to the original signal by itself, without any additional lowpass filtering, and in essence represents a possible, although admittedly very coarse, interpolation between the sample values. Alternatively, in some applications, we may wish to perform some smoother interpolation between sample values. In the next section, we explore in more detail the general concept of interpreting the reconstruction of a signal from its samples as a process of interpolation.

7.2 RECONSTRUCTION OF A SIGNAL FROM ITS SAMPLES USING INTERPOLATION

Interpolation, that is, the fitting of a continuous signal to a set of sample values, is a commonly used procedure for reconstructing a function, either approximately or exactly, from samples. One simple interpolation procedure is the zero-order hold discussed in Section 7.1. Another useful form of interpolation is *linear interpolation,* whereby adjacent sample points are connected by a straight line, as illustrated in Figure 7.9. In more

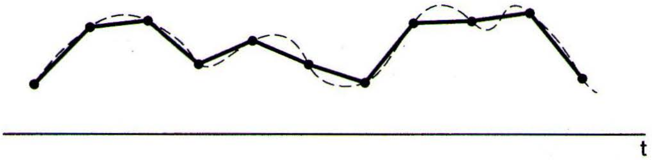


Figure 7.9 Linear interpolation between sample points. The dashed curve represents the original signal and the solid curve the linear interpolation.

complicated interpolation formulas, sample points may be connected by higher order polynomials or other mathematical functions.

As we have seen in Section 7.1, for a band-limited signal, if the sampling instants are sufficiently close, then the signal can be reconstructed exactly; i.e., through the use of a lowpass filter, exact interpolation can be carried out between the sample points. The interpretation of the reconstruction of $x(t)$ as a process of interpolation becomes evident when we consider the effect in the time domain of the lowpass filter in Figure 7.4. In particular, the output is

$$x_r(t) = x_p(t) * h(t)$$

or, with $x_p(t)$ given by eq. (7.3),

$$x_r(t) = \sum_{n=-\infty}^{+\infty} x(nT)h(t - nT). \tag{7.9}$$

Equation (7.9) describes how to fit a continuous curve between the sample points $x(nT)$ and consequently represents an interpolation formula. For the ideal lowpass filter $H(j\omega)$ in Figure 7.4,

$$h(t) = \frac{\omega_c T \sin(\omega_c t)}{\pi \omega_c t}, \tag{7.10}$$

so that

$$x_r(t) = \sum_{n=-\infty}^{+\infty} x(nT)\frac{\omega_c T}{\pi}\frac{\sin(\omega_c(t - nT))}{\omega_c(t - nT)}. \tag{7.11}$$

The reconstruction according to eq. (7.11) with $\omega_c = \omega_s/2$ is illustrated in Figure 7.10. Figure 7.10(a) represents the original band-limited signal $x(t)$, and Figure 7.10(b) represents $x_p(t)$, the impulse train of samples. In Figure 7.10(c), the superposition of the individual terms in eq. (7.11) is illustrated.

Interpolation using the impulse response of an ideal lowpass filter as in eq. (7.11) is commonly referred to as *band-limited interpolation,* since it implements exact reconstruction if $x(t)$ is band limited and the sampling frequency satisfies the conditions of the sampling theorem. As we have indicated, in many cases it is preferable to use a less accurate, but simpler, filter or, equivalently, a simpler interpolating function than the function in eq. (7.10). For example, the zero-order hold can be viewed as a form of interpolation between sample values in which the interpolating function $h(t)$ is the impulse response $h_0(t)$ depicted in Figure 7.6. In that sense, with $x_0(t)$ in the figure corresponding to the approximation to $x(t)$, the system $h_0(t)$ represents an approximation to the ideal lowpass filter required for the exact interpolation. Figure 7.11 shows the magnitude of the transfer function of the zero-order-hold interpolating filter, superimposed on the desired transfer function of the exact interpolating filter.

Both from Figure 7.11 and from Figure 7.6, we see that the zero-order hold is a very rough approximation, although in some cases it is sufficient. For example, if additional

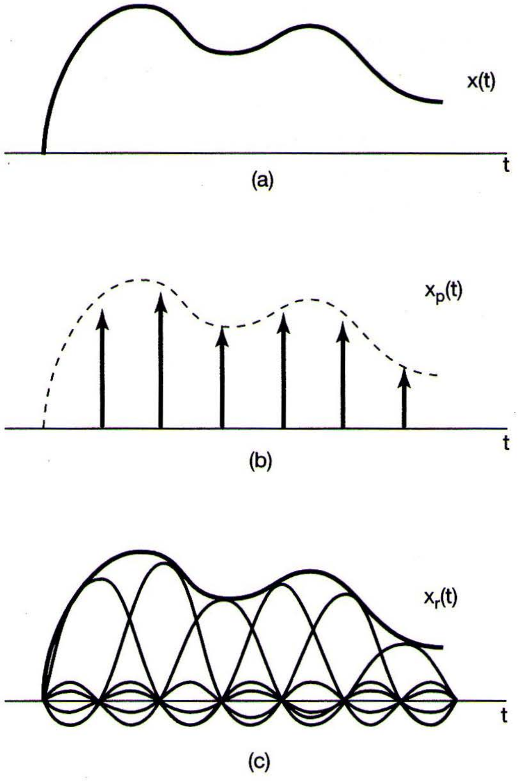


Figure 7.10 Ideal band-limited interpolation using the sinc function: (a) band-limited signal $x(t)$; (b) impulse train of samples of $x(t)$; (c) ideal band-limited interpolation in which the impulse train is replaced by a superposition of sinc functions [eq. (7.11)].

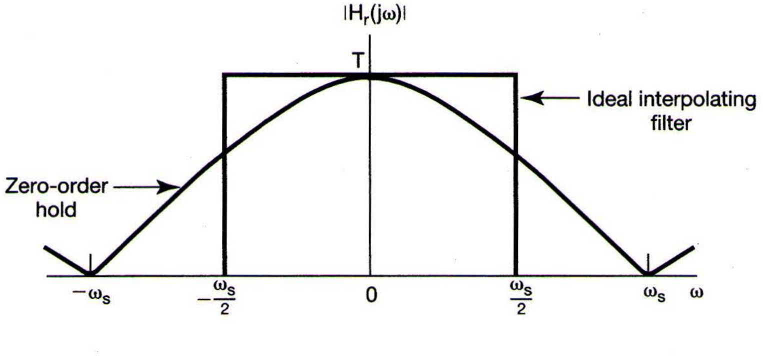
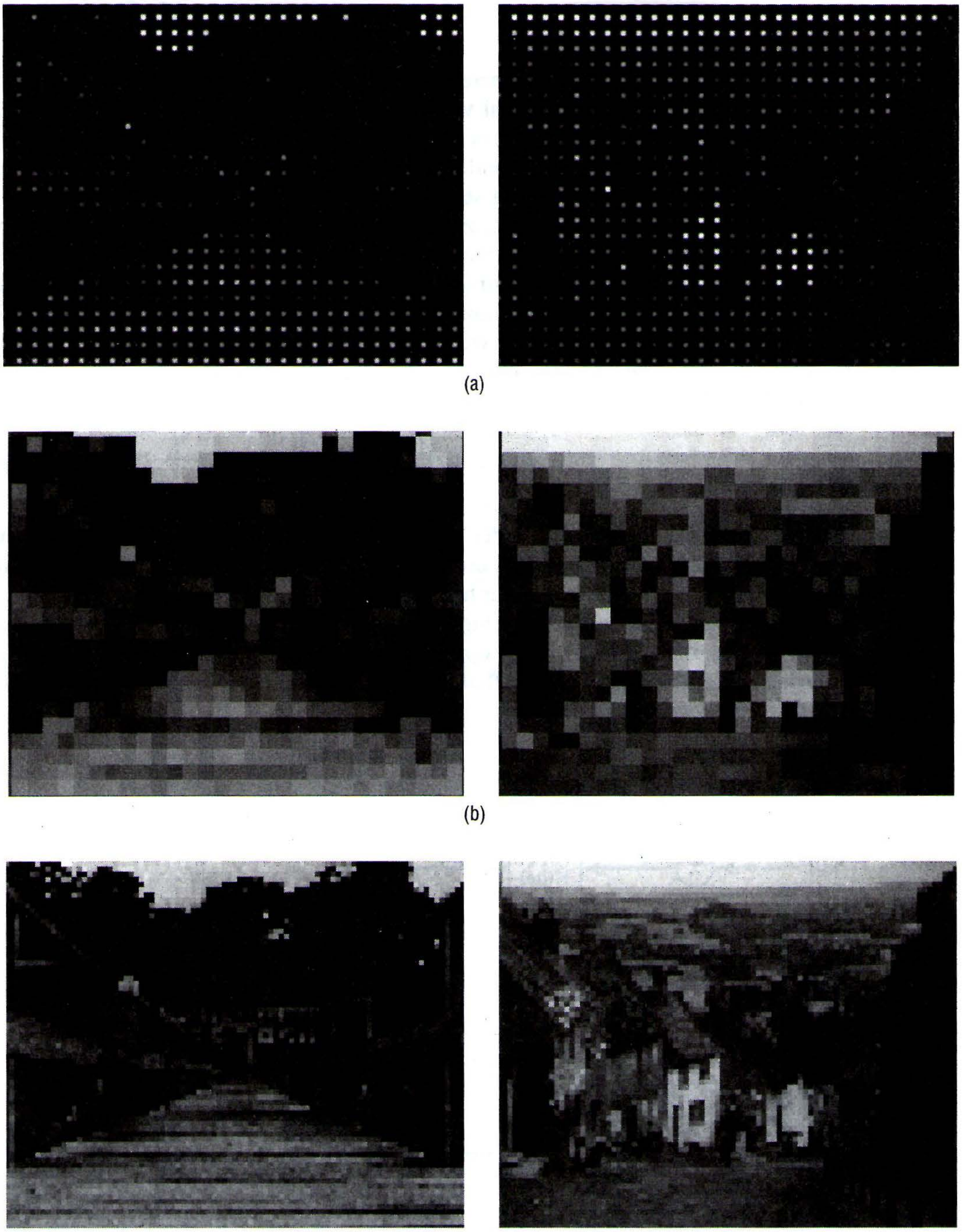


Figure 7.11 Transfer function for the zero-order hold and for the ideal interpolating filter.

lowpass filtering is naturally applied in a given application, it will tend to improve the overall interpolation. This is illustrated in the case of pictures in Figure 7.12. Figure 7.12(a) shows pictures with impulse sampling (i.e., sampling with spatially narrow pulses). Figure 7.12(b) is the result of applying a two-dimensional zero-order hold to Figure 7.12(a), with a resulting mosaic effect. However, the human visual system inherently imposes lowpass filtering, and consequently, when viewed at a distance, the discontinuities in the mosaic are smoothed. For example, in Figure 7.12(c) a



(a)

(b)

(c)

Figure 7.12 (a) The original pictures of Figures 6.2(a) and (g) with impulse sampling; (b) zero-order hold applied to the pictures in (a). The visual system naturally introduces lowpass filtering with a cutoff frequency that decreases with distance. Thus, when viewed at a distance, the discontinuities in the mosaic in Figure 7.12(b) are smoothed; (c) result of applying a zero-order hold after impulse sampling with one-fourth the horizontal and vertical spacing used in (a) and (b).

zero-order hold is again used, but here the sample spacing in each direction is one-fourth that in Figure 7.12(a). With normal viewing, considerable lowpass filtering is naturally applied, although the mosaic effect is still evident.

If the crude interpolation provided by the zero-order hold is insufficient, we can use a variety of smoother interpolation strategies, some of which are known collectively as *higher order holds.* In particular, the zero-order hold produces an output signal, as in Figure 7.5, that is discontinuous. In contrast, linear interpolation, as illustrated in Figure 7.9, yields reconstructions that are continuous, although with discontinous derivatives due to the changes in slope at the sample points. Linear interpolation, which is sometimes referred to as a first-order hold, can also be viewed as interpolation in the form of Figure 7.4 and eq. (7.9) with $h(t)$ triangular, as illustrated in Figure 7.13. The associated transfer function is also shown in the figure and is

$$H(j\omega) = \frac{1}{T}\left[\frac{\sin(\omega T/2)}{\omega/2}\right]^2. \tag{7.12}$$

The transfer function of the first-order hold is shown superimposed on the transfer function for the ideal interpolating filter. Figure 7.14 corresponds to the same pictures as those in Figure 7.12(b), but with a first-order hold applied to the sampled picture. In an analogous fashion, we can define second- and higher order holds that produce reconstructions with a higher degree of smoothness. For example, the output of a second-order hold provides an interpolation of the sample values that is continuous and has a continuous first derivative and discontinuous second derivative.

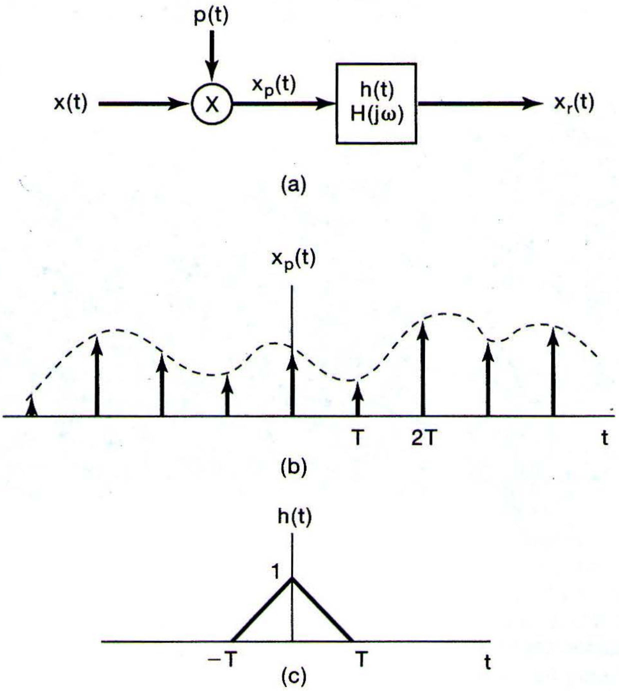
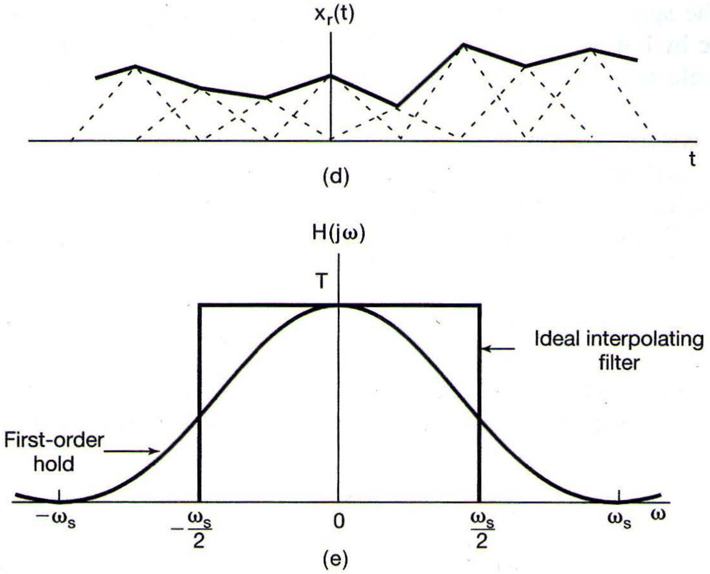


Figure 7.13 Linear interpolation (first-order hold) as impulse-train sampling followed by convolution with a triangular impulse response: (a) system for sampling and reconstruction; (b) impulse train of samples; (c) impulse response representing a first-order hold;

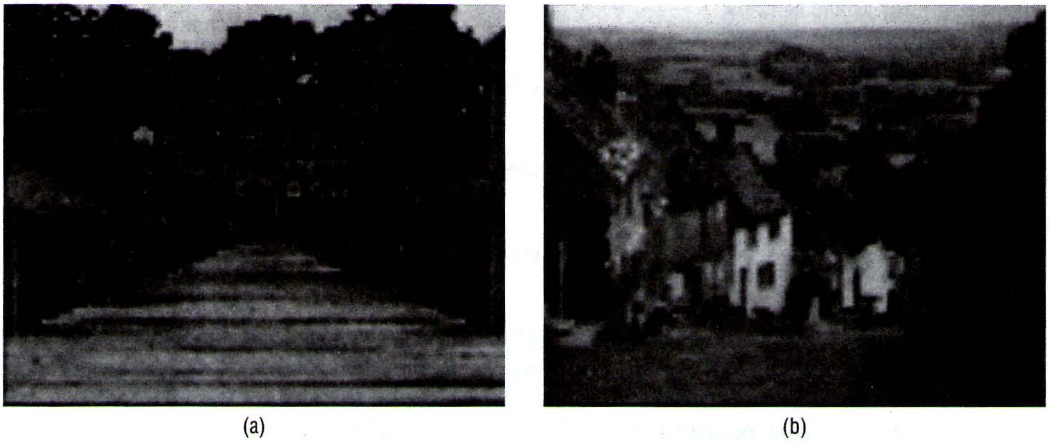


$x_r(t)$
t
(d)
$H(j\omega)$
T
Ideal interpolating filter
First-order hold
$-\omega_s$
$-\frac{\omega_s}{2}$
0
$\frac{\omega_s}{2}$
ω_s
ω
(e)

Figure 7.13 *Continued* (d) first-order hold applied to the sampled signal; (e) comparison of transfer function of ideal interpolating filter and first-order hold.

7.3 THE EFFECT OF UNDERSAMPLING: ALIASING

In previous sections in this chapter, it was assumed that the sampling frequency was sufficiently high that the conditions of the sampling theorem were met. As illustrated in Figure 7.3, with $\omega_s > 2\omega_M$, the spectrum of the sampled signal consists of scaled replications of the spectrum of $x(t)$, and this forms the basis for the sampling theorem. When


(a)

(b)

Figure 7.14 Result of applying a first-order hold rather than a zero-order hold after impulse sampling with one-third the horizontal and vertical spacing used in Figures 7.12(a) and (b).

$\omega_s < 2\omega_M$, $X(j\omega)$, the spectrum of $x(t)$, is no longer replicated in $X_p(j\omega)$ and thus is no longer recoverable by lowpass filtering. This effect, in which the individual terms in eq. (7.6) overlap, is referred to as *aliasing,* and in this section we explore its effect and consequences.

Clearly, if the system of Figure 7.4 is applied to a signal with $\omega_s < 2\omega_M$, the reconstructed signal $x_r(t)$ will no longer be equal to $x(t)$. However, as explored in Problem 7.25, the original signal and the signal $x_r(t)$ that is reconstructed using band-limited interpolation will always be equal at the sampling instants; that is, for any choice of ω_s,

$$x_r(nT) = x(nT), \qquad n = 0, \pm 1, \pm 2, \ldots. \tag{7.13}$$

Some insight into the relationship between $x(t)$ and $x_r(t)$ when $\omega_s < 2\omega_M$ is provided by considering in more detail the comparatively simple case of a sinusoidal signal. Thus, let

$$x(t) = \cos \omega_0 t, \tag{7.14}$$

with Fourier transform $X(j\omega)$ as indicated in Figure 7.15(a). In this figure, we have graphically distinguished the impulse at ω_0 from that at $-\omega_0$ for convenience. Let us consider $X_p(j\omega)$, the spectrum of the sampled signal, and focus in particular on the effect of a change in the frequency ω_0 with the sampling frequency ω_s fixed. In Figures 7.15(b)–(e), we illustrate $X_p(j\omega)$ for several values of ω_0. Also indicated by a dashed line is the passband of the lowpass filter of Figure 7.4 with $\omega_c = \omega_s/2$. Note that no aliasing occurs in (b) and (c), since $\omega_0 < \omega_s/2$, whereas aliasing does occur in (d) and (e). For each of the four cases, the lowpass filtered output $x_r(t)$ is given as follows:

(a) $\omega_0 = \dfrac{\omega_s}{6}; \qquad x_r(t) = \cos \omega_0 t = x(t)$

(b) $\omega_0 = \dfrac{2\omega_s}{6}; \qquad x_r(t) = \cos \omega_0 t = x(t)$

(c) $\omega_0 = \dfrac{4\omega_s}{6}; \qquad x_r(t) = \cos(\omega_s - \omega_0)t \neq x(t)$

(d) $\omega_0 = \dfrac{5\omega_s}{6}; \qquad x_r(t) = \cos(\omega_s - \omega_0)t \neq x(t).$

When aliasing occurs, the original frequency ω_0 takes on the identity of a lower frequency, $\omega_s - \omega_0$. For $\omega_s/2 < \omega_0 < \omega_s$, as ω_0 increases relative to ω_s, the output frequency $\omega_s - \omega_0$ decreases. When $\omega_s = \omega_0$, for example, the reconstructed signal is a constant. This is consistent with the fact that, when sampling once per cycle, the samples are all equal and would be identical to those obtained by sampling a constant signal ($\omega_0 = 0$). In Figure 7.16, we have depicted, for each of the four cases in Figure 7.15, the signal $x(t)$, its samples, and the reconstructed signal $x_r(t)$. From the figure, we can see how the lowpass

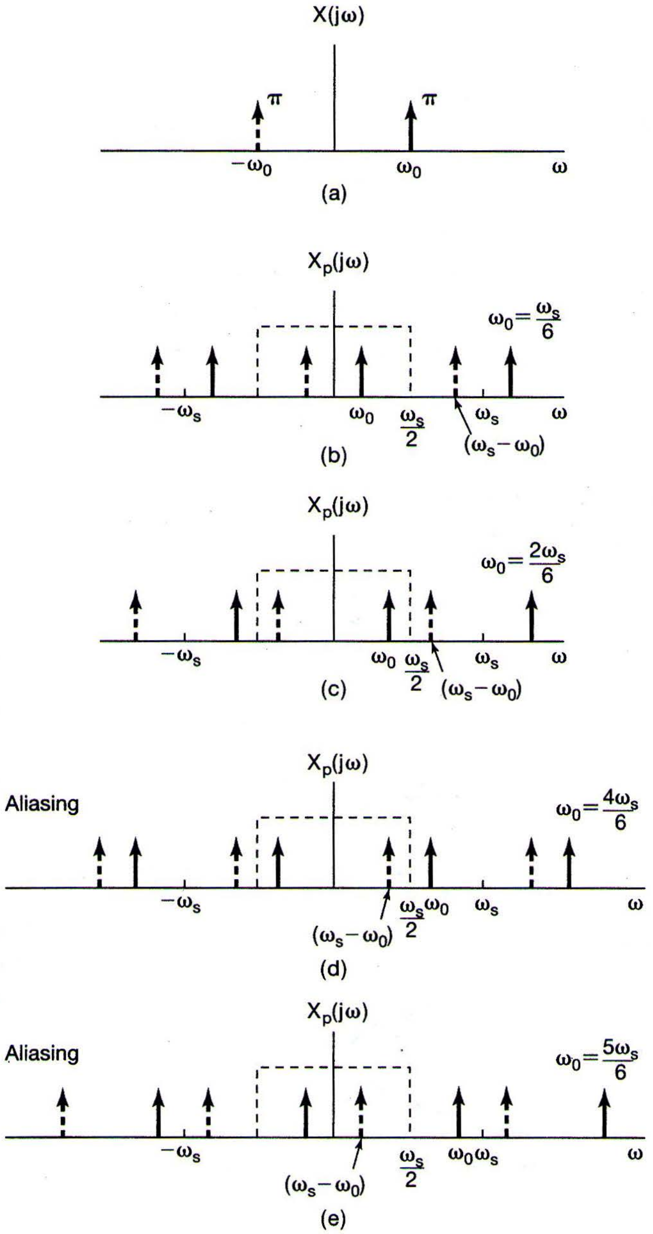

X(jω)
π
π
−ω0
ω0
ω
(a)
Xp(jω)
ω0 = ωs/6
−ωs
ω0
ωs/2
ωs
ω
(ωs − ω0)
(b)
Xp(jω)
ω0 = 2ωs/6
−ωs
ω0
ωs/2
ωs
ω
(ωs − ω0)
(c)
Aliasing
Xp(jω)
ω0 = 4ωs/6
−ωs
ωs/2
ω0
ωs
ω
(ωs − ω0)
(d)
Aliasing
Xp(jω)
ω0 = 5ωs/6
−ωs
ωs/2
ω0
ωs
ω
(ωs − ω0)
(e)

Figure 7.15 Effect in the frequency domain of oversampling and undersampling: (a) spectrum of original sinusoidal signal; (b), (c) spectrum of sampled signal with $\omega_s > 2\omega_0$; (d), (e) spectrum of sampled signal with $\omega_s < 2\omega_0$. As we increase ω_0 in moving from (b) through (d), the impulses drawn with solid lines move to the right, while the impulses drawn with dashed lines move to the left. In (d) and (e), these impulses have moved sufficiently that there is a change in the ones falling within the passband of the ideal lowpass filter.

filter interpolates between the samples, in particular always fitting a sinusoid of frequency less than $\omega_s/2$ to the samples of $x(t)$.

As a variation on the preceding examples, consider the signal

$$x(t) = \cos(\omega_0 t + \phi). \tag{7.15}$$

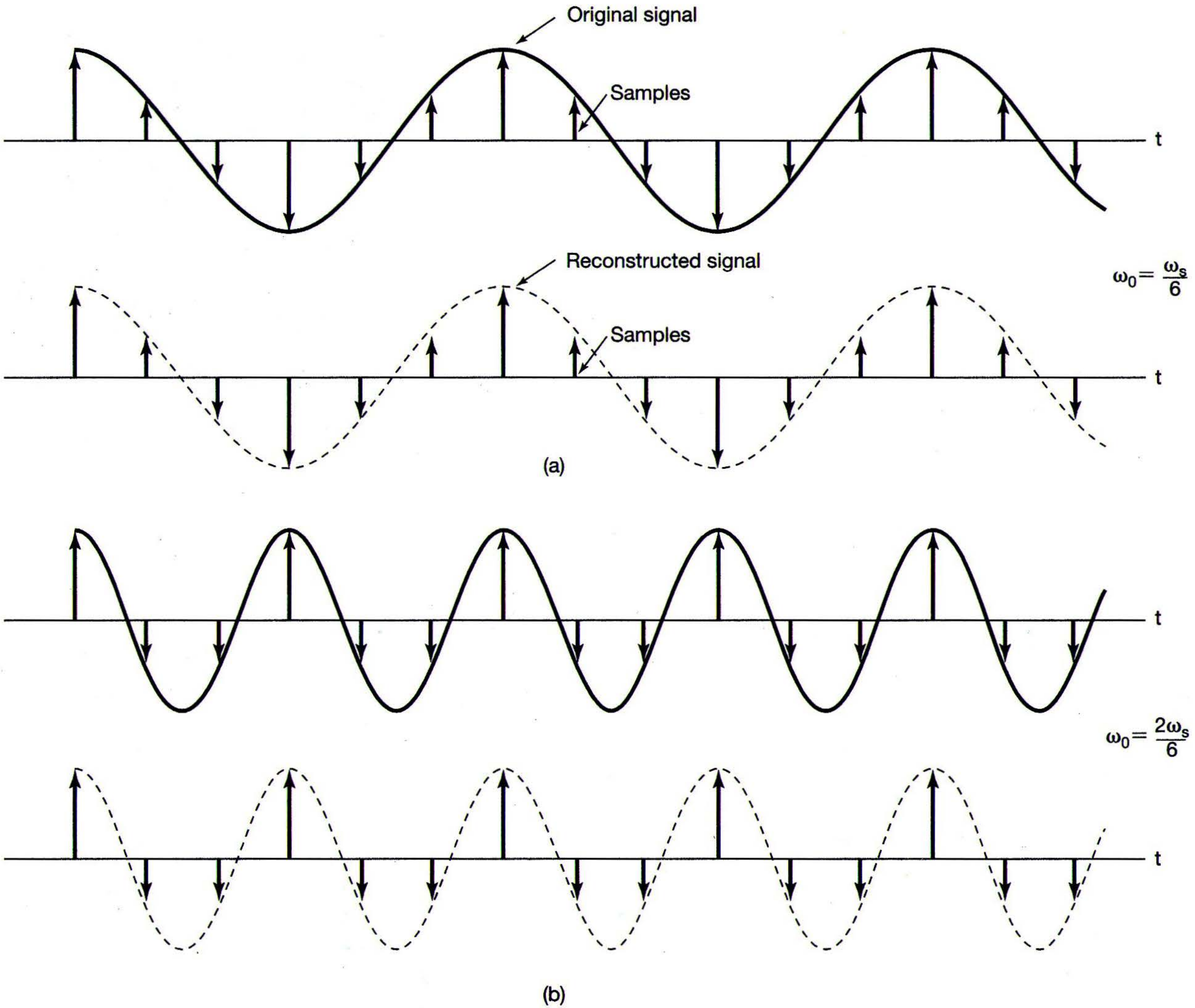


Figure 7.16 Effect of aliasing on a sinusoidal signal. For each of four values of ω_0, the original sinusoidal signal (solid curve), its samples, and the reconstructed signal (dashed curve) are illustrated: (a) $\omega_0 = \omega_s/6$; (b) $\omega_0 = 2\omega_s/6$; (c) $\omega_0 = 4\omega_s/6$; (d) $\omega_0 = 5\omega_s/6$. In (a) and (b) no aliasing occurs, whereas in (c) and (d) there is aliasing.

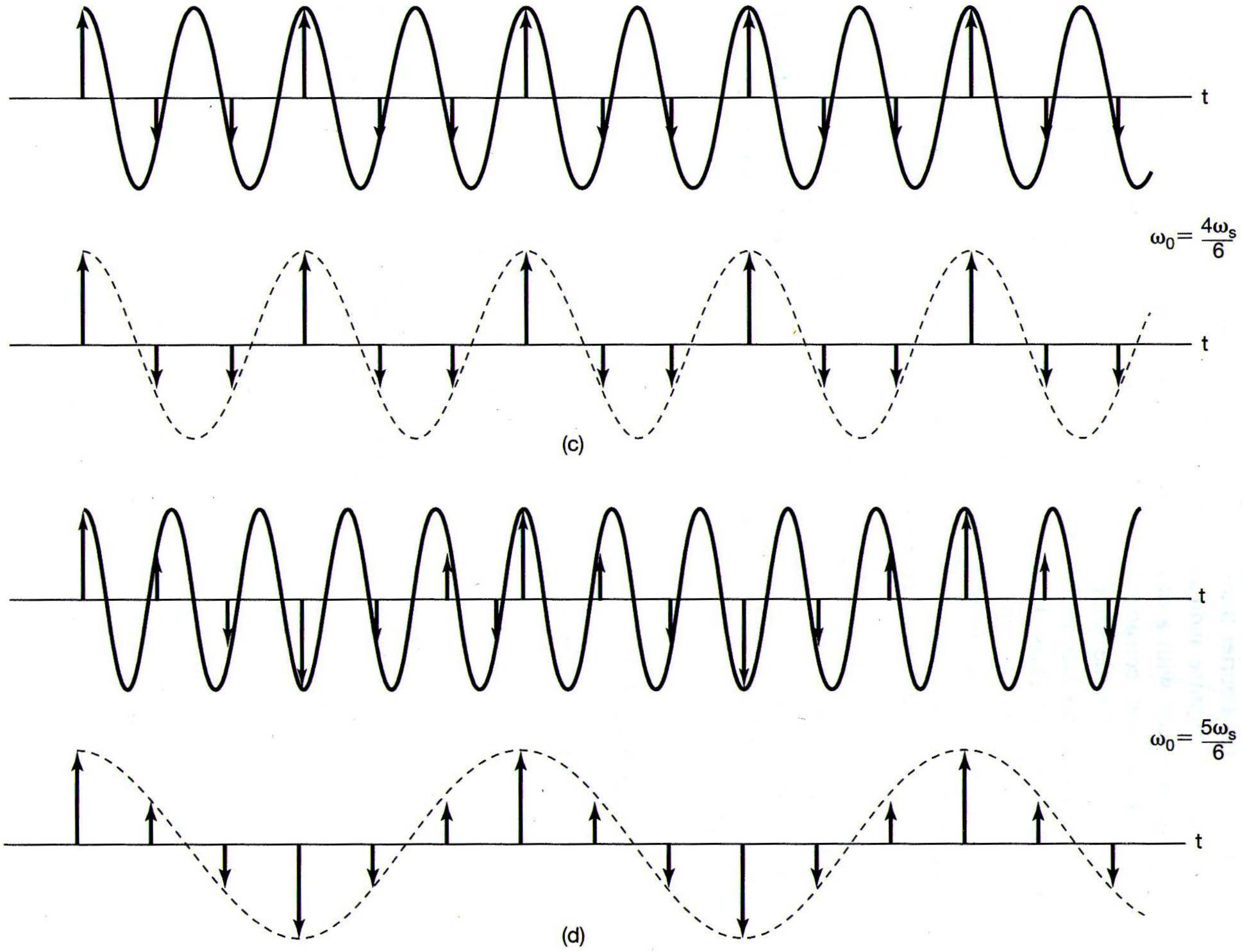

t
$\omega_0 = \frac{4\omega_s}{6}$
t
(c)
t
$\omega_0 = \frac{5\omega_s}{6}$
t
(d)

Figure 7.16 *Continued*

In this case, the Fourier transform of $x(t)$ is essentially the same as Figure 7.15(a), except that the impulse indicated with a solid line now has amplitude $\pi e^{j\phi}$, while the impulse indicated with a dashed line has amplitude with the opposite phase, namely, $\pi e^{-j\phi}$. If we now consider the same set of choices for ω_0 as in Figure 7.15, the resulting spectra for the sampled versions of $\cos(\omega_0 t + \phi)$ are exactly as in the figure, with all solid impulses having amplitude $\pi e^{j\phi}$ and all dashed ones having amplitude $\pi e^{-j\phi}$. Again, in cases (b) and (c) the condition of the sampling theorem is met, so that $x_r(t) = \cos(\omega_0 t + \phi) = x(t)$, while in (d) and (e) we again have aliasing. However, we now see that there has been a reversal in the solid and dashed impulses appearing in the passband of the lowpass filter. As a result, we find that in these cases, $x_r(t) = \cos[(\omega_s - \omega_0)t - \phi]$, where we have a change in the sign of the phase ϕ, i.e., a *phase reversal.*

It is important to note that the sampling theorem explicity requires that the sampling frequency be *greater than* twice the highest frequency in the signal, rather than greater than or equal to twice the highest frequency. The next example illustrates that sampling a sinusoidal signal at exactly twice its frequency (i.e., exactly two samples per cycle) is not sufficient.

Example 7.1

Consider the sinusoidal signal

$$x(t) = \cos\left(\frac{\omega_s}{2}t + \phi\right),$$

and suppose that this signal is sampled, using impulse sampling, at exactly twice the frequency of the sinusoid, i.e., at sampling frequency ω_s. As shown in Problem 7.39, if this impulse-sampled signal is applied as the input to an ideal lowpass filter with cutoff frequency $\omega_s/2$, the resulting output is

$$x_r(t) = (\cos\phi)\cos\left(\frac{\omega_s}{2}t\right).$$

As a consequence, we see that perfect reconstruction of $x(t)$ occurs only in the case in which the phase ϕ is zero (or an integer multiple of 2π). Otherwise, the signal $x_r(t)$ does not equal $x(t)$.

As an extreme example, consider the case in which $\phi = -\pi/2$, so that

$$x(t) = \sin\left(\frac{\omega_s}{2}t\right).$$

This signal is sketched in Figure 7.17. We observe that the values of the signal at integer multiples of the sampling period $2\pi/\omega_s$ are zero. Consequently, sampling at this rate produces a signal that is identically zero, and when this zero input is applied to the ideal lowpass filter, the resulting output $x_r(t)$ is also identically zero.

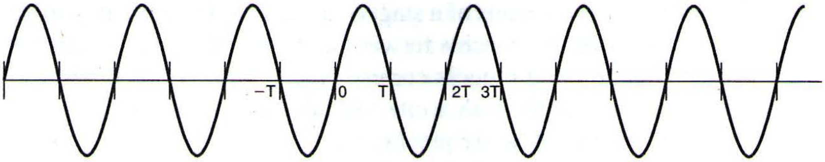


Figure 7.17 Sinusoidal signal for Example 7.1.

The effect of undersampling, whereby higher frequencies are reflected into lower frequencies, is the principle on which the stroboscopic effect is based. Consider, for example, the situation depicted in Figure 7.18, in which we have a disc rotating at a constant rate with a single radial line marked on the disc. The flashing strobe acts as a sampling system, since it illuminates the disc for extremely brief time intervals at a periodic rate. When the strobe frequency is much higher than the rotational speed of the disc, the speed of rotation of the disc is perceived correctly. When the strobe frequency becomes less than twice the rotational frequency of the disc, the rotation appears to be at a lower frequency than is actually the case. Furthermore, because of phase reversal, the disc will appear to be rotating in the wrong direction! Roughly speaking, if we track the position of a fixed line on the disc at successive samples, then when $\omega_0 < \omega_s < 2\omega_0$, so that we sample somewhat more frequently than once per revolution, samples of the disc will show the fixed line in positions that are successively displaced in a counterclockwise direction, opposite to the clockwise rotation of the disc itself. At one flash per revolution, corresponding to $\omega_s = \omega_0$, the radial line appears stationary (i.e., the rotational frequency of the disc and its harmonics have been aliased to zero frequency). A similar effect is commonly observed in Western movies,

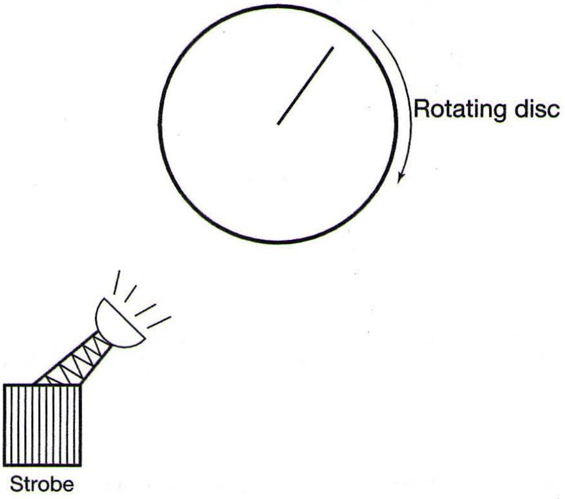


Figure 7.18 Strobe effect.

The transformation of $x_c(t)$ to $x_d[n]$ corresponding to the first system in Figure 7.19 will be referred to as *continuous-to-discrete-time conversion* and will be abbreviated C/D. The reverse operation corresponding to the third system in Figure 7.19 will be abbreviated D/C, representing *discrete-time to continuous-time conversion*. The D/C operation performs an interpolation between the sample values provided to it as input. That is, the D/C operation produces a continuous-time signal $y_c(t)$ which is related to the discrete-time signal $y_d[n]$ by

$$y_d[n] = y_c(nT).$$

This notation is made explicit in Figure 7.20. In systems such as digital computers and digital systems for which the discrete-time signal is represented in digital form, the device commonly used to implement the C/D conversion is referred to as an *analog-to-digital* (A-to-D) converter, and the device used to implement the D/C conversion is referred to as a *digital-to-analog* (D-to-A) converter.

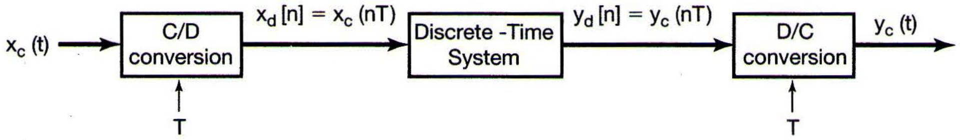


Figure 7.20 Notation for continuous-to-discrete-time conversion and discrete-to-continuous-time conversion. T represents the sampling period.

To understand further the relationship between the continuous-time signal $x_c(t)$ and its discrete-time representation $x_d[n]$, it is helpful to represent C/D as a process of periodic sampling followed by a mapping of the impulse train to a sequence. These two steps are illustrated in Figure 7.21. In the first step, representing the sampling process, the impulse train $x_p(t)$ corresponds to a sequence of impulses with amplitudes corresponding to the samples of $x_c(t)$ and with a time spacing equal to the sampling period T. In the conversion from the impulse train to the discrete-time sequence, we obtain $x_d[n]$, corresponding to the same sequence of samples of $x_c(t)$, but with unity spacing in terms of the new independent variable n. Thus, in effect, the conversion from the impulse train sequence of samples to the discrete-time sequence of samples can be thought of as a normalization in time. This normalization in converting $x_p(t)$ to $x_d[n]$ is evident in Figures 7.21(b) and (c), in which $x_p(t)$ and $x_d[n]$ are respectively illustrated for sampling rates of $T = T_1$ and $T = 2T_1$.

It is also instructive to examine the processing stages in Figure 7.19 in the frequency domain. Since we will be dealing with Fourier transforms in both continuous and discrete time, in this section only we distinguish the continuous-time and discrete-time frequency variables by using ω in continuous time and Ω in discrete time. For example, the continuous-time Fourier transforms of $x_c(t)$ and $y_c(t)$ are $X_c(j\omega)$ and $Y_c(j\omega)$, respectively, while the discrete-time Fourier transforms of $x_d[n]$ and $y_d[n]$ are $X_d(e^{j\Omega})$ and $Y_d(e^{j\Omega})$, respectively.

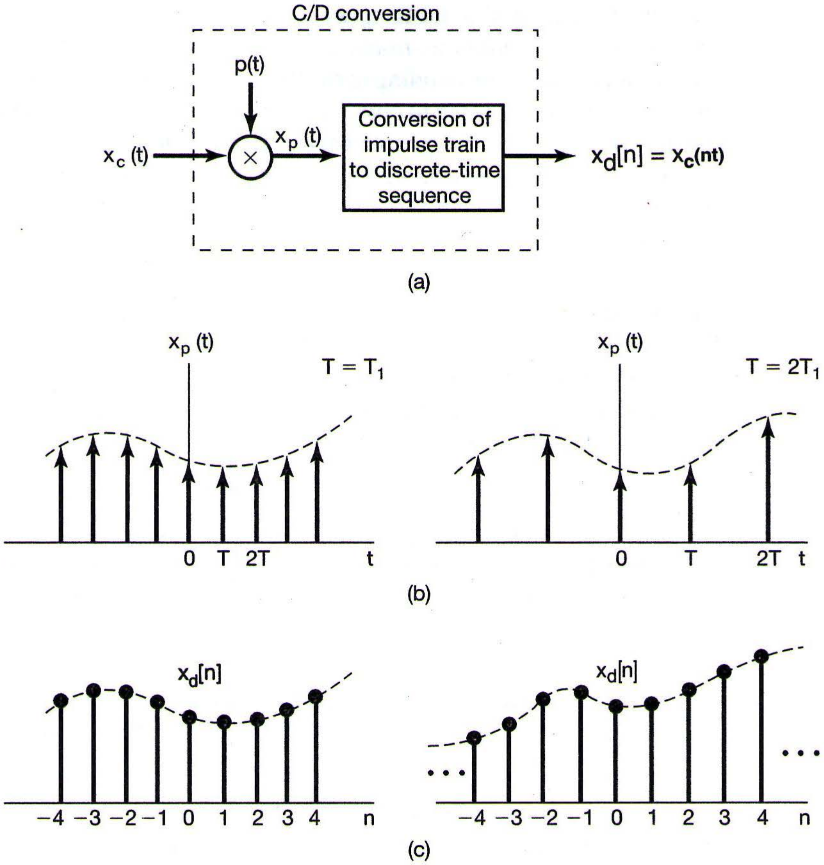

C/D conversion
p(t)
x_c (t)
x_p (t)
Conversion of impulse train to discrete-time sequence
x_d[n] = x_c(nt)
(a)
x_p (t)
T = T_1
0 T 2T t
x_p (t)
T = 2T_1
0 T 2T t
(b)
x_d[n]
−4 −3 −2 −1 0 1 2 3 4 n
x_d[n]
−4 −3 −2 −1 0 1 2 3 4 n
(c)

Figure 7.21 Sampling with a periodic impulse train followed by conversion to a discrete-time sequence: (a) overall system; (b) $x_p(t)$ for two sampling rates. The dashed envelope represents $x_c(t)$; (c) the output sequence for the two different sampling rates.

To begin let us express $X_p(j\omega)$, the continuous-time Fourier transform of $x_p(t)$, in terms of the sample values of $x_c(t)$ by applying the Fourier transform to eq. (7.3). Since

$$x_p(t) = \sum_{n=-\infty}^{+\infty} x_c(nT)\delta(t - nT), \tag{7.17}$$

and since the transform of $\delta(t - nT)$ is $e^{-j\omega nT}$, it follows that

$$X_p(j\omega) = \sum_{n=-\infty}^{+\infty} x_c(nT)e^{-j\omega nT} \tag{7.18}$$

Now consider the discrete-time Fourier transform of $x_d[n]$, that is,

$$X_d(e^{j\Omega}) = \sum_{n=-\infty}^{+\infty} x_d[n]e^{-j\Omega n}, \tag{7.19}$$

or, using eq. (7.16),

$$X_d(e^{j\Omega}) = \sum_{n=-\infty}^{+\infty} x_c(nT)e^{-j\Omega n}. \tag{7.20}$$

Comparing eqs. (7.18) and (7.20), we see that $X_d(e^{j\Omega})$ and $X_p(j\omega)$ are related through

$$X_d(e^{j\Omega}) = X_p(j\Omega/T). \tag{7.21}$$

Also, recall that, as developed in eq. (7.6) and illustrated in Figure 7.3,

$$X_p(j\omega) = \frac{1}{T}\sum_{k=-\infty}^{+\infty} X_c(j(\omega - k\omega_s)). \tag{7.22}$$

Consequently,

$$X_d(e^{j\Omega}) = \frac{1}{T}\sum_{k=-\infty}^{+\infty} X_c(j(\Omega - 2\pi k)/T). \tag{7.23}$$

The relationship among $X_c(j\omega)$, $X_p(j\omega)$, and $X_d(e^{j\Omega})$ is illustrated in Figure 7.22 for two different sampling rates. Here, $X_d(e^{j\Omega})$ is a frequency-scaled version of $X_p(j\omega)$

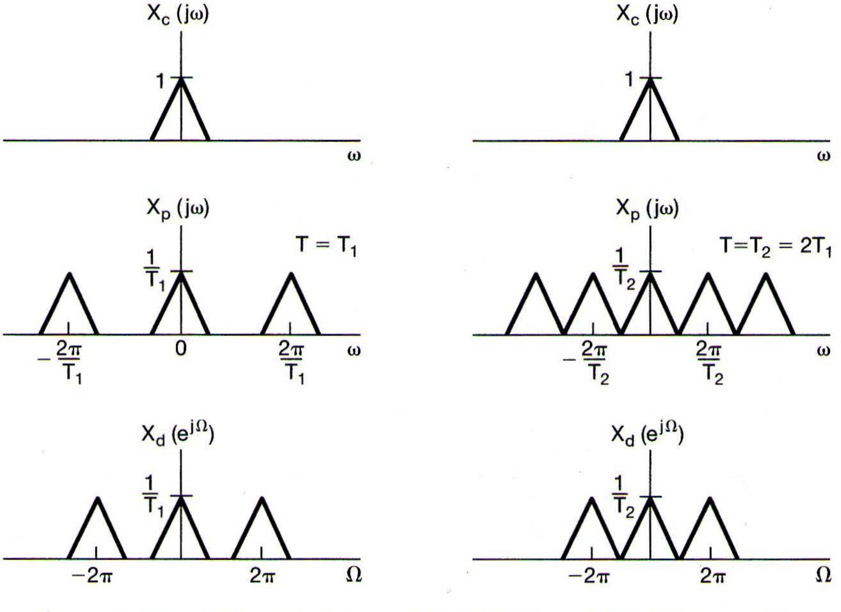

$X_c(j\omega)$
1
ω
$X_c(j\omega)$
1
ω
$X_p(j\omega)$
$\frac{1}{T_1}$
$T = T_1$
$-\frac{2\pi}{T_1}$
0
$\frac{2\pi}{T_1}$
ω
$X_p(j\omega)$
$\frac{1}{T_2}$
$T=T_2 = 2T_1$
$-\frac{2\pi}{T_2}$
$\frac{2\pi}{T_2}$
ω
$X_d(e^{j\Omega})$
$\frac{1}{T_1}$
-2π
2π
Ω
$X_d(e^{j\Omega})$
$\frac{1}{T_2}$
-2π
2π
Ω

Figure 7.22 Relationship between $X_c(j\omega)$, $X_p(j\omega)$, and $X_d(e^{j\Omega})$ for two different sampling rates.

and, in particular, is periodic in Ω with period 2π. This periodicity is, of course, characteristic of any discrete-time Fourier transform. The spectrum of $x_d[n]$ is related to that of $x_c(t)$ through periodic replication, represented by eq. (7.22), followed by linear frequency scaling, represented by eq. (7.21). The periodic replication is a consequence of the first step in the conversion process in Figure 7.21, namely, the impulse-train sampling. The linear frequency scaling in eq. (7.21) can be thought of informally as a consequence of the normalization in time introduced by converting from the impulse train $x_p(t)$ to the discrete-time sequence $x_d[n]$. From the time-scaling property of the Fourier transform in Section 4.3.5, scaling of the time axis by $1/T$ will introduce a scaling of the frequency axis by T. Thus, the relationship $\Omega = \omega T$ is consistent with the notion that, in converting from $x_p(t)$ to $x_d[n]$, the time axis is scaled by $1/T$.

In the overall system of Figure 7.19, after processing with a discrete-time system, the resulting sequence is converted back to a continuous-time signal. This process is the reverse of the steps in Figure 7.21. Specifically, from the sequence $y_d[n]$, a continuous-time impulse train $y_p(t)$ can be generated. Recovery of the continuous-time signal $y_c(t)$ from this impulse train is then accomplished by means of lowpass filtering, as illustrated in Figure 7.23.

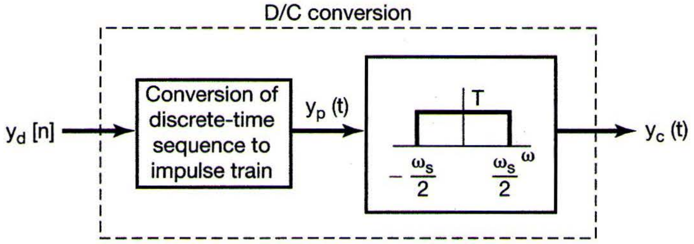

D/C conversion
$y_d[n]$
Conversion of discrete-time sequence to impulse train
$y_p(t)$
T
$-\frac{\omega_s}{2}$
$\frac{\omega_s}{2}$
ω
$y_c(t)$

Figure 7.23 Conversion of a discrete-time sequence to a continuous-time signal.

Now let us consider the overall system of Figure 7.19, represented as shown in Figure 7.24. Clearly, if the discrete-time system is an identity system (i.e., $x_d[n] = y_d[n]$), then, assuming that the conditions of the sampling theorem are met, the overall system will be an identity system. The characteristics of the overall system with a more general frequency response $H_d(e^{j\Omega})$ are perhaps best understood by examining the representative example depicted in Figure 7.25. On the left-hand side of the figure are the representative

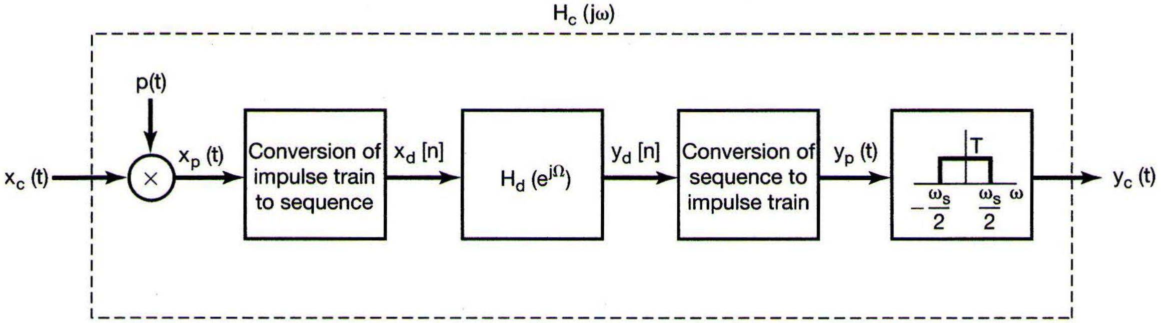

$H_c(j\omega)$
p(t)
$x_c(t)$
×
$x_p(t)$
Conversion of impulse train to sequence
$x_d[n]$
$H_d(e^{j\Omega})$
$y_d[n]$
Conversion of sequence to impulse train
$y_p(t)$
T
$-\frac{\omega_s}{2}$
$\frac{\omega_s}{2}$
ω
$y_c(t)$

Figure 7.24 Overall system for filtering a continuous-time signal using a discrete-time filter.

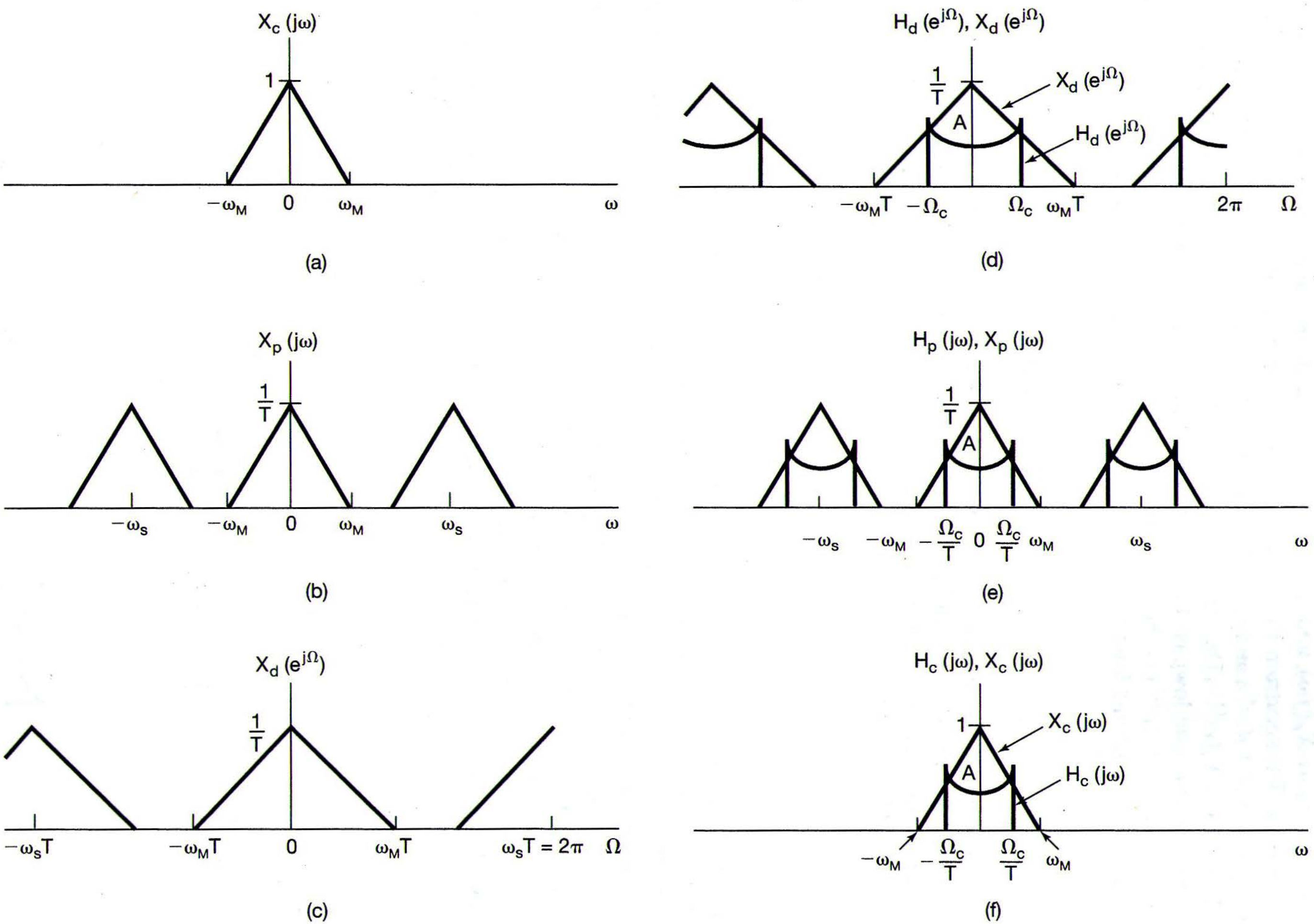


Figure 7.25 Frequency-domain illustration of the system of Figure 7.24: (a) continuous-time spectrum $X_c(j\omega)$; (b) spectrum after impulse-train sampling; (c) spectrum of discrete-time sequence $x_d[n]$; (d) $H_d(e^{j\Omega})$ and $X_d(e^{j\Omega})$ that are multiplied to form $Y_d(e^{j\Omega})$; (e) spectra that are multiplied to form $Y_p(j\omega)$; (f) spectra that are multiplied to form $Y_c(j\omega)$.

spectra $X_c(j\omega)$, $X_p(j\omega)$, and $X_d(e^{j\Omega})$, where we assume that $\omega_M < \omega_s/2$, so that there is no aliasing. The spectrum $Y_d(e^{j\Omega})$ corresponding to the output of the discrete-time filter is the product of $X_d(e^{j\Omega})$ and $H_d(e^{j\Omega})$, and this is depicted in Figure 7.25(d) by overlaying $H_d(e^{j\Omega})$ and $X_d(e^{j\Omega})$. The transformation to $Y_c(j\omega)$ then corresponds to applying a frequency scaling and lowpass filtering, resulting in the spectra indicated in Figure 7.25(e) and (f). Since $Y_d(e^{j\Omega})$ is the product of the two overlaid spectra in Figure 7.25(d), the frequency scaling and lowpass filtering are applied to both. In comparing Figures 7.25(a) and (f), we see that

$$Y_c(j\omega) = X_c(j\omega)H_d(e^{j\omega T}). \tag{7.24}$$

Consequently, for inputs that are sufficiently band limited, so that the sampling theorem is satisfied, the overall system of Figure 7.24 is, in fact, equivalent to a continuous-time LTI system with frequency response $H_c(j\omega)$ which is related to the discrete-time frequency response $H_d(e^{j\Omega})$ through

$$H_c(j\omega) = \begin{cases} H_d(e^{j\omega T}), & |\omega| < \omega_s/2 \\ 0, & |\omega| > \omega_s/2 \end{cases}. \tag{7.25}$$

The equivalent frequency response for this continuous-time filter is one period of the frequency response of the discrete-time filter with a linear scale change applied to the frequency axis. This relationship between the discrete-time frequency response and the equivalent continuous-time frequency response is illustrated in Figure 7.26.

The equivalence of the overall system of Figure 7.24 to an LTI system is somewhat surprising in view of the fact that multiplication by an impulse train is *not* a time-invariant operation. In fact, the overall system of Figure 7.24 is not time invariant for arbitrary inputs. For example, if $x_c(t)$ was a narrow rectangular pulse of duration less than T, then a time shift of $x_c(t)$ could generate a sequence $x[n]$ that either had all zero values or had one nonzero value, depending on the alignment of the rectangular pulse relative to the

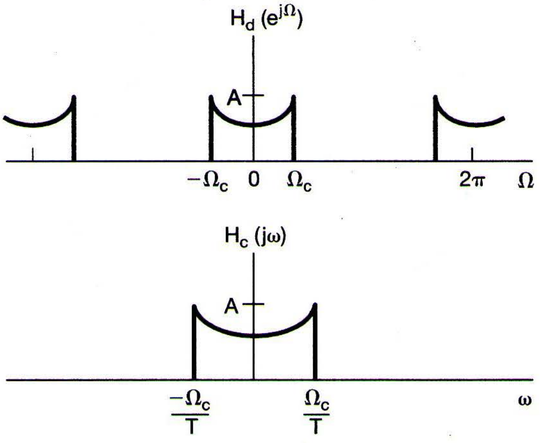


Figure 7.26 Discrete-time frequency response and the equivalent continuous-time frequency response for the system of Figure 7.24.

sampling impulse train. However, as suggested by the spectra of Figure 7.25, for *band-limited input signals* with a sampling rate sufficiently high so as to avoid aliasing, the system of Figure 7.24 *is* equivalent to a continuous-time LTI system. For such inputs, Figure 7.24 and eq. (7.25) provide the conceptual basis for continuous-time processing using discrete-time filters. This is now explored further in the context of some examples.

7.4.1 Digital Differentiator

Consider the discrete-time implementation of a continuous-time band-limited differentiating filter. As discussed in Section 3.9.1, the frequency response of a continuous-time differentiating filter is

$$H_c(j\omega) = j\omega, \tag{7.26}$$

and that of a band-limited differentiator with cutoff frequency ω_c is

$$H_c(j\omega) = \begin{cases} j\omega, & |\omega| < \omega_c \\ 0, & |\omega| > \omega_c \end{cases}, \tag{7.27}$$

as sketched in Figure 7.27. Using eq. (7.25) with a sampling frequency $\omega_s = 2\omega_c$, we see that the corresponding discrete-time transfer function is

$$H_d(e^{j\Omega}) = j\left(\frac{\Omega}{T}\right), \quad |\Omega| < \pi, \tag{7.28}$$

as sketched in Figure 7.28. With this discrete-time transfer function, $y_c(t)$ in Figure 7.24 will be the derivative of $x_c(t)$, assuming that there is no aliasing in sampling $x_c(t)$.

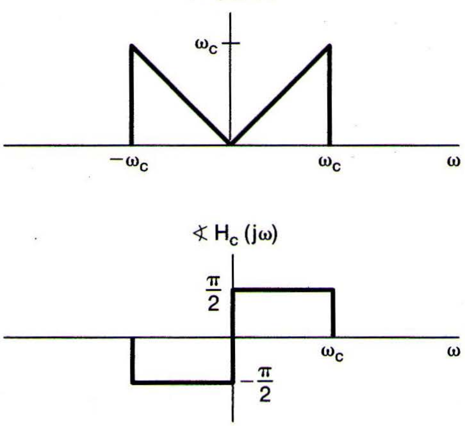


Figure 7.27 Frequency response of a continuous-time ideal band-limited differentiator $H_c(j\omega) = j\omega$, $|\omega| < \omega_c$.

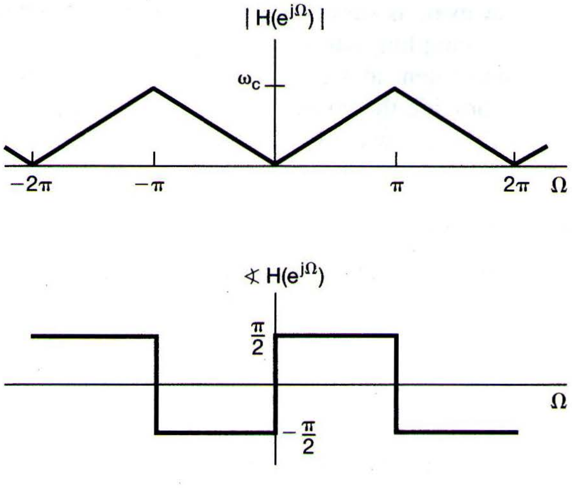


Figure 7.28 Frequency response of discrete-time filter used to implement a continuous-time band-limited differentiator.

Example 7.2

By considering the output of the digital differentiator for a continuous-time sinc input, we may conveniently determine the impulse response $h_d[n]$ of the discrete-time filter in the implementation of the digital differentiator. With reference to Figure 7.24, let

$$x_c(t) = \frac{\sin(\pi t/T)}{\pi t}, \tag{7.29}$$

where T is the sampling period. Then

$$X_c(j\omega) = \begin{cases} 1, & |\omega| < \pi/T \\ 0, & \text{otherwise} \end{cases},$$

which is sufficiently band limited to ensure that sampling $x_c(t)$ at frequency $\omega_s = 2\pi/T$ does not give rise to any aliasing. It follows that the output of the digital differentiator is

$$y_c(t) = \frac{d}{dt}x_c(t) = \frac{\cos(\pi t/T)}{Tt} - \frac{\sin(\pi t/T)}{\pi t^2}. \tag{7.30}$$

For $x_c(t)$ as given by eq. (7.29), the corresponding signal $x_d[n]$ in Figure 7.24 may be expressed as

$$x_d[n] = x_c(nT) = \frac{1}{T}\delta[n]. \tag{7.31}$$

That is, for $n \neq 0$, $x_c(nT) = 0$, while

$$x_d[0] = x_c(0) = \frac{1}{T}$$

which can be verified by l'Hôpital's rule. We can similarly evaluate $y_d[n]$ in Figure 7.24 corresponding to $y_c(t)$ in eq. (7.30). Specifically

$$y_d[n] = y_c(nT) = \begin{cases} \dfrac{(-1)^n}{nT^2}, & n \neq 0 \\ 0 & n = 0 \end{cases}. \tag{7.32}$$

which can be verified for $n \neq 0$ by direct substitution into eq. (7.30) and for $n = 0$ by application of l'Hôpital's rule.

Thus when the input to the discrete-time filter given by eq. (7.28) is the scaled unit impulse in eq. (7.31), the resulting output is given by eq. (7.32). We then conclude that the impulse response of this filter is given by

$$h_d[n] = \begin{cases} \dfrac{(-1)^n}{nT}, & n \neq 0 \\ 0, & n = 0 \end{cases}$$

7.4.2 Half-Sample Delay

In this section, we consider the implementation of a time shift (delay) of a continuous-time signal through the use of a system in the form of Figure 7.19. Thus, we require that the input and output of the overall system be related by

$$y_c(t) = x_c(t - \Delta) \tag{7.33}$$

when the input $x_c(t)$ is band limited and the sampling rate is high enough to avoid aliasing and where Δ represents the delay time. From the time-shifting property derived in Section 4.3.2

$$Y_c(j\omega) = e^{-j\omega\Delta} X_c(j\omega).$$

From eq. (7.25), the equivalent continuous-time system to be implemented must be band limited. Therefore, we take

$$H_c(j\omega) = \begin{cases} e^{-j\omega\Delta}, & |\omega| < \omega_c \\ 0, & \text{otherwise} \end{cases}, \tag{7.34}$$

where ω_c is the cutoff frequency of the continuous-time filter. That is, $H_c(j\omega)$ corresponds to a time shift as in eq. (7.33) for band-limited signals and rejects all frequencies greater than ω_c. The magnitude and phase of the frequency response are shown in Figure 7.29(a). With the sampling frequency ω_s taken as $\omega_s = 2\omega_c$, the corresponding discrete-time

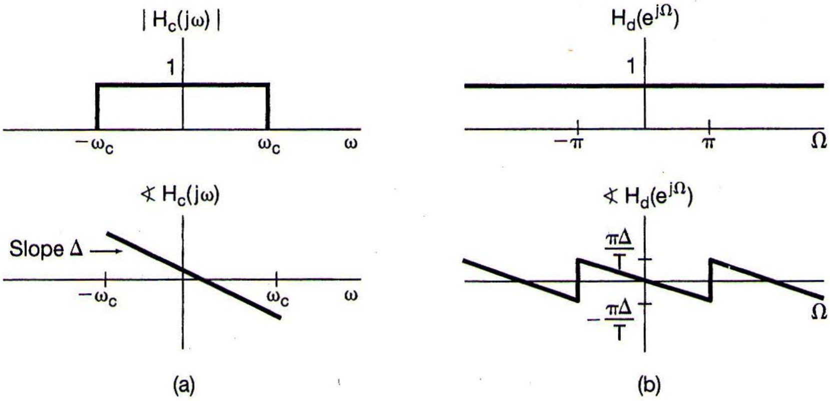


Figure 7.29 (a) Magnitude and phase of the frequency response for a continuous-time delay; (b) magnitude and phase of the frequency response for the corresponding discrete-time delay.

frequency response is

$$H_d(e^{j\Omega}) = e^{-j\Omega\Delta/T}, \quad |\Omega| < \pi, \tag{7.35}$$

and is shown in Figure 7.29(b).

For appropriately band-limited inputs, the output of the system of Figure 7.24 with $H_d(e^{j\Omega})$ as in eq. (7.35) is a delayed replica of the input. For Δ/T an integer, the sequence $y_d[n]$ is a delayed replica of $x_d[n]$; that is,

$$y_d[n] = x_d\left[n - \frac{\Delta}{T}\right]. \tag{7.36}$$

For Δ/T not an integer, eq. (7.36), as written, has no meaning, since sequences are defined only at integer values of the index. However, we can interpret the relationship between $x_d[n]$ and $y_d[n]$ in these cases in terms of band-limited interpolation. The signals $x_c(t)$ and $x_d[n]$ are related through sampling and band-limited interpolation, as are $y_c(t)$ and $y_d[n]$. With $H_d(e^{j\Omega})$ in eq. (7.35), $y_d[n]$ is equal to samples of a shifted version of the band-limited interpolation of the sequence $x_d[n]$. This is illustrated in Figure 7.30 with $\Delta/T = 1/2$, which is sometimes referred to as a half-sample delay.

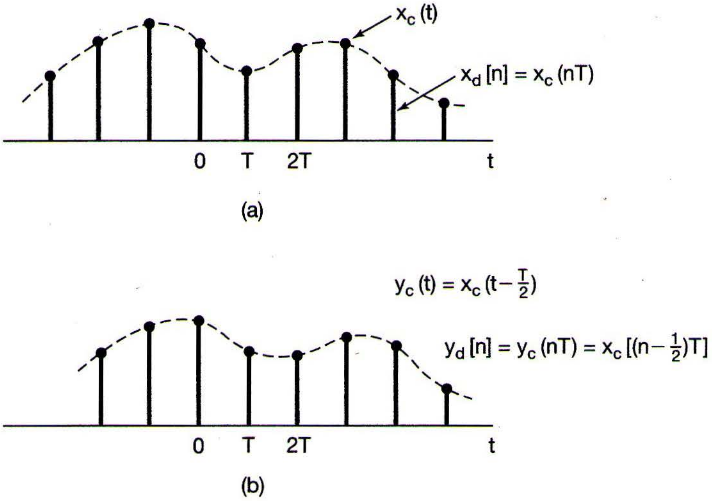

$x_c(t)$
$x_d[n] = x_c(nT)$
0
T
2T
t
(a)
$y_c(t) = x_c(t - \frac{T}{2})$
$y_d[n] = y_c(nT) = x_c[(n-\frac{1}{2})T]$
0
T
2T
t
(b)

Figure 7.30 (a) Sequence of samples of a continuous-time signal $x_c(t)$; (b) sequence in (a) with a half-sample delay.

Example 7.3

The approach in Example 7.2 is also applicable to determining the impulse response $h_d[n]$ of the discrete-time filter in the half-sample delay system. With reference to Figure 7.24, let

$$x_c(t) = \frac{\sin(\pi t/T)}{\pi t}. \tag{7.37}$$

It follows from Example 7.2 that

$$x_d[n] = x_c(nT) = \frac{1}{T}\delta[n].$$

Also, since there is no aliasing for the band-limited input in eq. (7.37), the output of the half-sample delay system is

$$y_c(t) = x_c(t - T/2) = \frac{\sin(\pi(t - T/2)/T)}{\pi(t - T/2)},$$

and the sequence $y_d[n]$ in Figure 7.24 is

$$y_d[n] = y_c(nT) = \frac{\sin(\pi(n - \frac{1}{2}))}{T\pi(n - \frac{1}{2})}.$$

We conclude that

$$h[n] = \frac{\sin(\pi(n - \frac{1}{2}))}{\pi(n - \frac{1}{2})}.$$

7.5 SAMPLING OF DISCRETE-TIME SIGNALS

Thus far in this chapter, we have considered the sampling of continuous-time signals, and in addition to developing the analysis necessary to understand continuous-time sampling, we have introduced a number of its applications. As we will see in this section, a very similar set of properties and results with a number of important applications can be developed for sampling of discrete-time signals.

7.5.1 Impulse-Train Sampling

In analogy with continuous-time sampling as carried out using the system of Figure 7.2, sampling of a discrete-time signal can be represented as shown in Figure 7.31. Here, the new sequence $x_p[n]$ resulting from the sampling process is equal to the original sequence $x[n]$ at integer multiples of the sampling period N and is zero at the intermediate samples; that is,

$$x_p[n] = \begin{cases} x[n], & \text{if } n = \text{an integer multiple of } N \\ 0, & \text{otherwise} \end{cases}. \tag{7.38}$$

As with continuous-time sampling in Section 7.1, the effect in the frequency domain of discrete-time sampling is seen by using the multiplication property developed in Section 5.5. Thus, with

$$x_p[n] = x[n]p[n] = \sum_{k=-\infty}^{+\infty} x[kN]\delta[n - kN], \tag{7.39}$$

we have, in the frequency domain,

$$X_p(e^{j\omega}) = \frac{1}{2\pi}\int_{2\pi} P(e^{j\theta})X(e^{j(\omega-\theta)})d\theta. \tag{7.40}$$

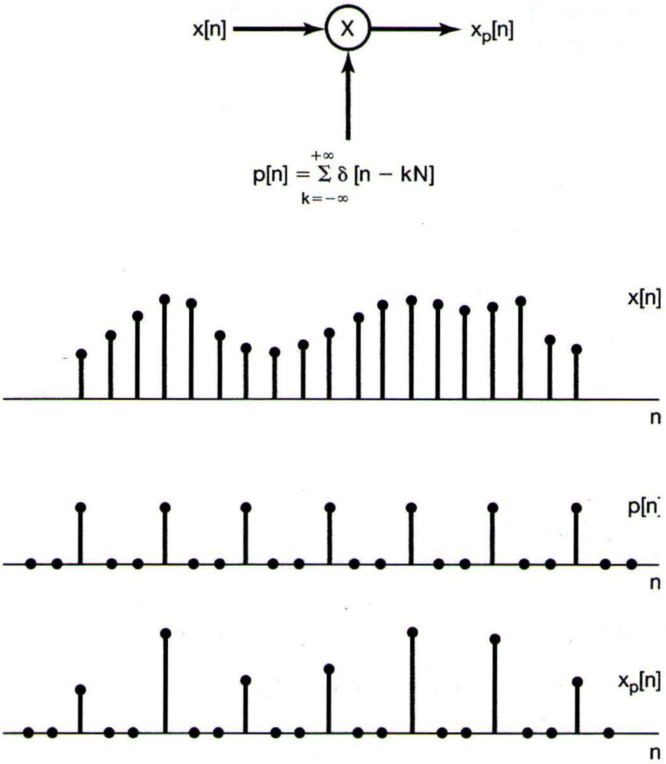


Figure 7.31 Discrete-time sampling.

As in Example 5.6, the Fourier transform of the sampling sequence $p[n]$ is

$$P(e^{j\omega}) = \frac{2\pi}{N} \sum_{k=-\infty}^{+\infty} \delta(\omega - k\omega_s), \tag{7.41}$$

where ω_s, the sampling frequency, equals $2\pi/N$. Combining eqs. (7.40) and (7.41), we have

$$X_p(e^{j\omega}) = \frac{1}{N} \sum_{k=0}^{N-1} X\left(e^{j(\omega - k\omega_s)}\right). \tag{7.42}$$

Equation (7.42) is the counterpart for discrete-time sampling of eq. (7.6) for continuous-time sampling and is illustrated in Figure 7.32. In Figure 7.32(c), with $\omega_s - \omega_M > \omega_M$, or equivalently, $\omega_s > 2\omega_M$, there is no aliasing [i.e., the nonzero portions of the replicas of $X(e^{j\omega})$ do not overlap], whereas with $\omega_s < 2\omega_M$, as in Figure 7.32(d), frequency-domain aliasing results. In the absence of aliasing, $X(e^{j\omega})$ is faithfully reproduced around $\omega = 0$ and integer multiples of 2π. Consequently, $x[n]$ can be recovered from $x_p[n]$ by means of a lowpass filter with gain N and a cutoff frequency greater than

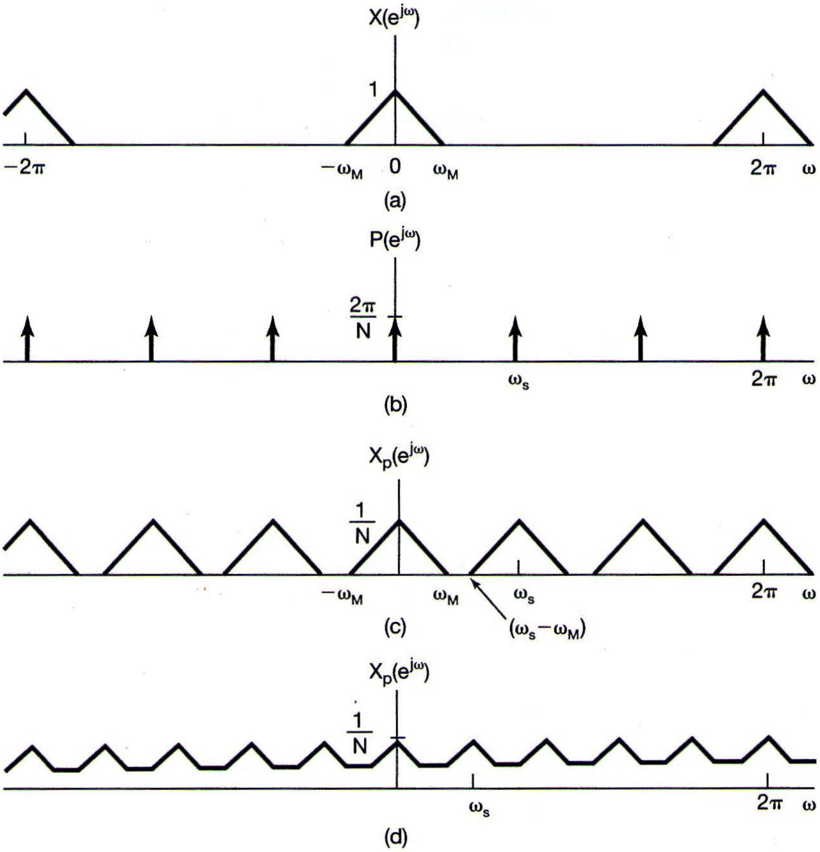


Figure 7.32 Effect in the frequency domain of impulse-train sampling of a discrete-time signal: (a) spectrum of original signal; (b) spectrum of sampling sequence; (c) spectrum of sampled signal with $\omega_s > 2\omega_M$; (d) spectrum of sampled signal with $\omega_s < 2\omega_M$. Note that aliasing occurs.

ω_M and less than $\omega_s - \omega_M$, as illustrated in Figure 7.33, where we have specified the cutoff frequency of the lowpass filter as $\omega_s/2$. If the overall system of Figure 7.33(a) is applied to a sequence for which $\omega_s < 2\omega_M$, so that aliasing results, $x_r[n]$ will no longer be equal to $x[n]$. However, as with continuous-time sampling, the two sequences *will* be equal at multiples of the sampling period; that is, corresponding to eq. (7.13), we have

$$x_r[kN] = x[kN], \quad k = 0, \pm 1, \pm 2, \ldots, \tag{7.43}$$

independently of whether aliasing occurs. (See Problem 7.46.)

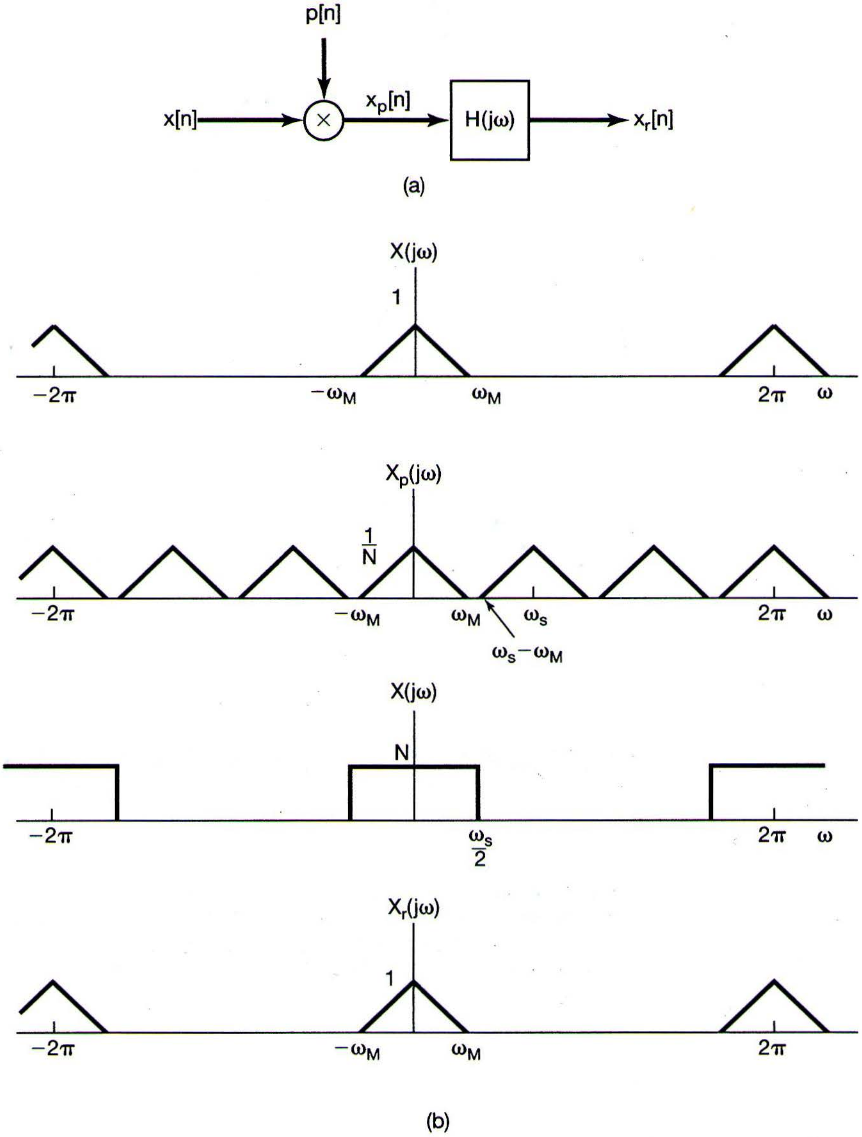

p[n]
x[n]
x_p[n]
H(jω)
x_r[n]
(a)
X(jω)
1
−2π
−ω_M
ω_M
2π
ω
X_p(jω)
1/N
−2π
−ω_M
ω_M
ω_s
ω_s−ω_M
2π
ω
X(jω)
N
−2π
ω_s/2
2π
ω
X_r(jω)
1
−2π
−ω_M
ω_M
2π
(b)

Figure 7.33 Exact recovery of a discrete-time signal from its samples using an ideal lowpass filter: (a) block diagram for sampling and reconstruction of a band-limited signal from its samples; (b) spectrum of the signal $x[n]$; (c) spectrum of $x_p[n]$; (d) frequency response of an ideal lowpass filter with cutoff frequency $\omega_s/2$; (e) spectrum of the reconstructed signal $x_r[n]$. For the example depicted here $\omega_s > 2\omega_M$ so that no aliasing occurs and consequently $x_r[n] = x[n]$.

Example 7.4

Consider a sequence $x[n]$ whose Fourier transform $X(e^{j\omega})$ has the property that

$$X(e^{j\omega}) = 0 \qquad \text{for} \quad 2\pi/9 \leq |\omega| \leq \pi.$$

To determine the lowest rate at which $x[n]$ may be sampled without the possibility of aliasing, we must find the largest N such that

$$\frac{2\pi}{N} \geq 2\left(\frac{2\pi}{9}\right) \Longrightarrow N \leq 9/2.$$

We conclude that $N_{\max} = 4$, and the corresponding sampling frequency is $2\pi/4 = \pi/2$.

The reconstruction of $x[n]$ through the use of a lowpass filter applied to $x_p[n]$ can be interpreted in the time domain as an interpolation formula similar to eq. (7.11). With $h[n]$ denoting the impulse response of the lowpass filter, we have

$$h[n] = \frac{N\omega_c}{\pi} \frac{\sin \omega_c n}{\omega_c n}. \tag{7.44}$$

The reconstructed sequence is then

$$x_r[n] = x_p[n] * h[n], \tag{7.45}$$

or equivalently,

$$x_r[n] = \sum_{k=-\infty}^{+\infty} x[kN] \frac{N\omega_c}{\pi} \frac{\sin \omega_c(n - kN)}{\omega_c(n - kN)}. \tag{7.46}$$

Equation (7.46) represents ideal band-limited interpolation and requires the implementation of an ideal lowpass filter. In typical applications a suitable approximation for the lowpass filter in Figure 7.33 is used, in which case the equivalent interpolation formula is of the form

$$x_r[n] = \sum_{k=-\infty}^{+\infty} x[kN] h_r[n - kN], \tag{7.47}$$

where $h_r[n]$ is the impulse response of the interpolating filter. Some specific examples, including the discrete-time counterparts of the zero-order hold and first-order hold discussed in Section 7.2 for continuous-time interpolation, are considered in Problem 7.50.

7.5.2 Discrete-Time Decimation and Interpolation

There are a variety of important applications of the principles of discrete-time sampling, such as in filter design and implementation or in communication applications. In many of these applications it is inefficient to represent, transmit, or store the sampled sequence $x_p[n]$ directly in the form depicted in Figure 7.31, since, in between the sampling instants, $x_p[n]$ is *known* to be zero. Thus, the sampled sequence is typically replaced by a new sequence $x_b[n]$, which is simply every Nth value of $x_p[n]$; that is,

$$x_b[n] = x_p[nN]. \tag{7.48}$$

Also, equivalently,

$$x_b[n] = x[nN], \tag{7.49}$$

since $x_p[n]$ and $x[n]$ are equal at integer multiples of N. The operation of extracting every Nth sample is commonly referred to as *decimation*.[3] The relationship between $x[n]$, $x_p[n]$, and $x_b[n]$ is illustrated in Figure 7.34.

To determine the effect in the frequency domain of decimation, we wish to determine the relationship between $X_b(e^{j\omega})$—the Fourier transform of $x_b[n]$—and $X(e^{j\omega})$. To this end, we note that

$$X_b(e^{j\omega}) = \sum_{k=-\infty}^{+\infty} x_b[k]e^{-j\omega k}, \tag{7.50}$$

or, using eq. (7.48),

$$X_b(e^{j\omega}) = \sum_{k=-\infty}^{+\infty} x_p[kN]e^{-j\omega k}. \tag{7.51}$$

If we let $n = kN$, or equivalently $k = n/N$, we can write

$$X_b(e^{j\omega}) = \sum_{\substack{n=\text{integer}\\ \text{multiple of } N}} x_p[n]e^{-j\omega n/N},$$

and since $x_p[n] = 0$ when n is not an integer multiple of N, we can also write

$$X_b(e^{j\omega}) = \sum_{n=-\infty}^{+\infty} x_p[n]e^{-j\omega n/N}. \tag{7.52}$$

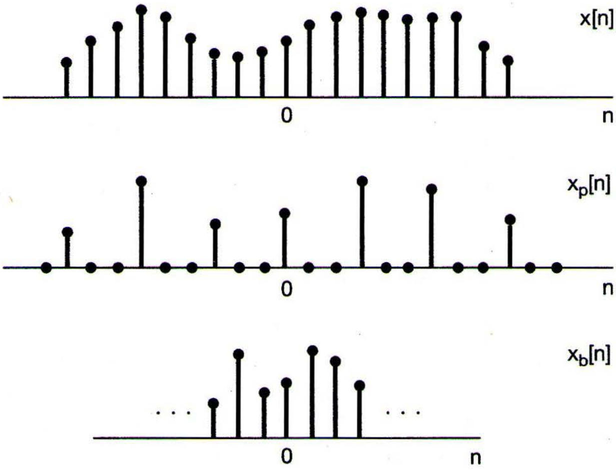


Figure 7.34 Relationship between $x_p[n]$ corresponding to sampling and $x_b[n]$ corresponding to decimation.

---

[3] Technically, decimation would correspond to extracting every *tenth* sample. However, it has become common terminology to refer to the operation as decimation even when N is not equal to 10.

Furthermore, we recognize the right-hand side of eq. (7.52) as the Fourier transform of $x_p[n]$; that is,

$$\sum_{n=-\infty}^{+\infty} x_p[n]e^{-j\omega n/N} = X_p(e^{j\omega/N}). \tag{7.53}$$

Thus, from eqs. (7.52) and (7.53), we conclude that

$$X_b(e^{j\omega}) = X_p(e^{j\omega/N}). \tag{7.54}$$

This relationship is illustrated in Figure 7.35, and from it, we observe that the spectra for the sampled sequence and the decimated sequence differ only in a frequency scaling or normalization. If the original spectrum $X(e^{j\omega})$ is appropriately band limited, so that there is no aliasing present in $X_p(e^{j\omega})$, then, as shown in the figure, the effect of decimation is to spread the spectrum of the original sequence over a larger portion of the frequency band.

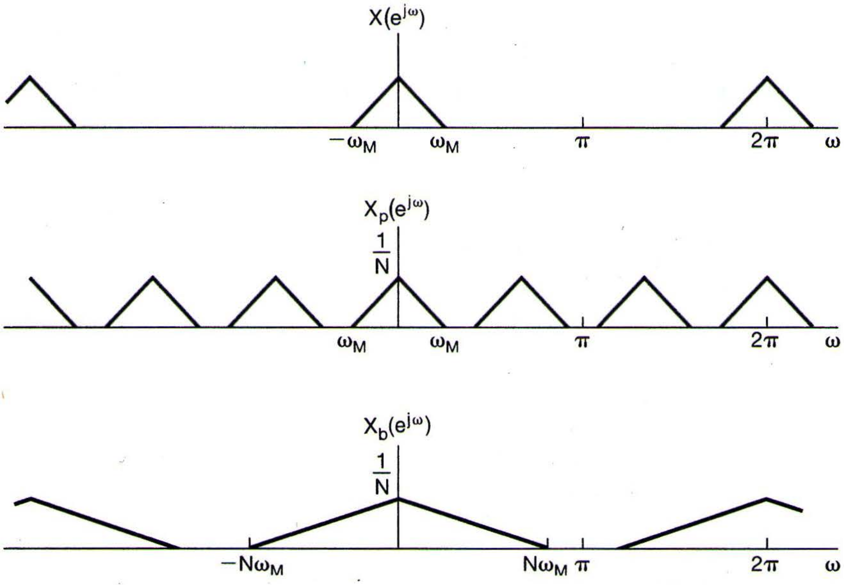


Figure 7.35 Frequency-domain illustration of the relationship between sampling and decimation.

If the original sequence $x[n]$ is obtained by sampling a continuous-time signal, the process of decimation can be viewed as reducing the sampling rate on the signal by a factor of N. To avoid aliasing, $X(e^{j\omega})$ cannot occupy the full frequency band. In other words, if the signal can be decimated without introducing aliasing, then the original continuous-time signal was oversampled, and thus, the sampling rate can be reduced without aliasing. With the interpretation of the sequence $x[n]$ as samples of a continuous-time signal, the process of decimation is often referred to as *downsampling*.

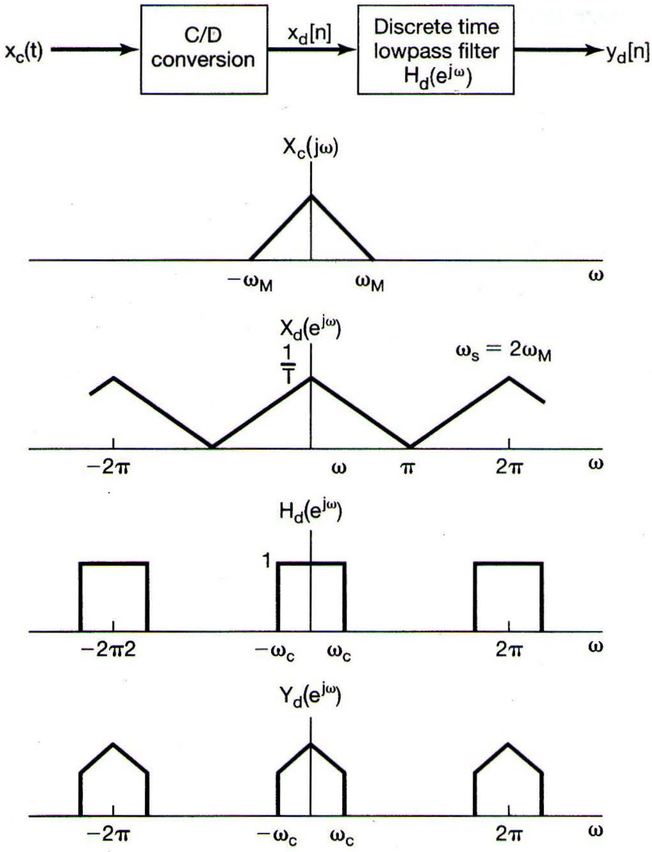

$x_c(t)$
C/D conversion
$x_d[n]$
Discrete time lowpass filter $H_d(e^{j\omega})$
$y_d[n]$
$X_c(j\omega)$
$-\omega_M$
ω_M
ω
$X_d(e^{j\omega})$
$\frac{1}{T}$
$\omega_s = 2\omega_M$
-2π
ω
π
2π
ω
$H_d(e^{j\omega})$
1
$-2\pi 2$
$-\omega_c$
ω_c
2π
ω
$Y_d(e^{j\omega})$
-2π
$-\omega_c$
ω_c
2π
ω

Figure 7.36 Continuous-time signal that was originally sampled at the Nyquist rate. After discrete-time filtering, the resulting sequence can be further downsampled. Here $X_c(j\omega)$ is the continuous-time Fourier transform of $x_c(t)$, $X_d(e^{j\omega})$ and $Y_d(e^{j\omega})$ are the discrete-time Fourier transforms of $x_d[n]$ and $y_d[n]$ respectively, and $H_d(e^{j\omega})$ is the frequency response of the discrete-time lowpass filter depicted in the block diagram.

In some applications in which a sequence is obtained by sampling a continuous-time signal, the original sampling rate may be as low as possible without introducing aliasing, but after additional processing and filtering, the bandwidth of the sequence may be reduced. An example of such a situation is shown in Figure 7.36. Since the output of the discrete-time filter is band limited, downsampling or decimation can be applied.

Just as in some applications it is useful to downsample, there are situations in which it is useful to convert a sequence to a *higher* equivalent sampling rate, a process referred to as *upsampling* or *interpolation.* Upsampling is basically the reverse of decimation or downsampling. As illustrated in Figures 7.34 and 7.35, in decimation we first sample and then retain only the sequence values at the sampling instants. To upsample, we reverse the process. For example, referring to Figure 7.34, we consider upsampling the sequence $x_b[n]$ to obtain $x[n]$. From $x_b[n]$, we form the sequence $x_p[n]$ by inserting $N-1$ points with zero amplitude between each of the values in $x_b[n]$. The interpolated sequence $x[n]$ is then obtained from $x_p[n]$ by lowpass filtering. The overall procedure is summarized in Figure 7.37.

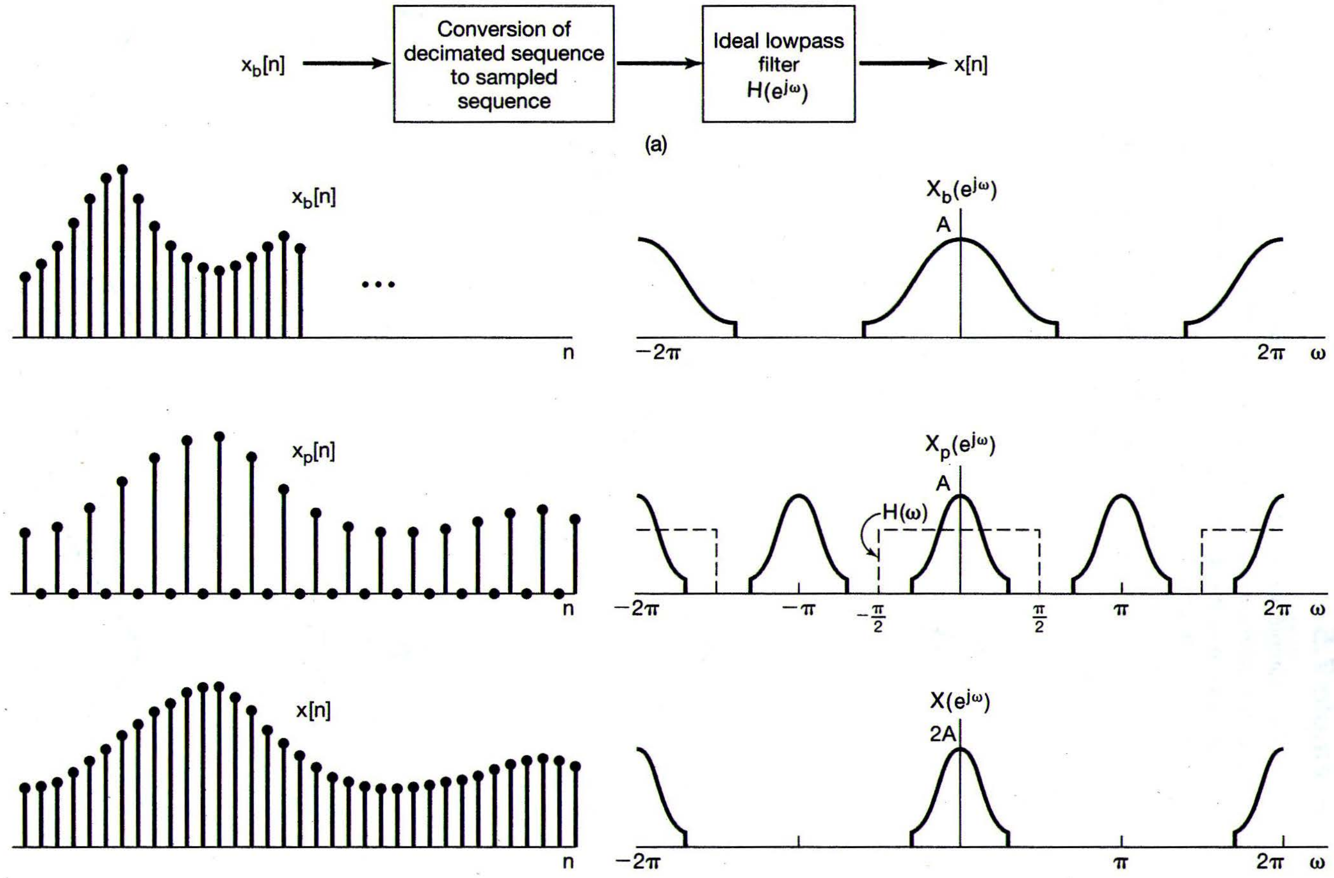

$x_b[n]$
Conversion of decimated sequence to sampled sequence
Ideal lowpass filter $H(e^{j\omega})$
$x[n]$
(a)
$x_b[n]$
n
$X_b(e^{j\omega})$
A
-2π
2π
ω
$x_p[n]$
n
$X_p(e^{j\omega})$
A
$H(\omega)$
-2π
$-\pi$
$-\frac{\pi}{2}$
$\frac{\pi}{2}$
π
2π
ω
$x[n]$
n
$X(e^{j\omega})$
2A
-2π
π
2π
ω

Figure 7.37 Upsampling: (a) overall system; (b) associated sequences and spectra for upsampling by a factor of 2.

Example 7.5

In this example, we illustrate how a combination of interpolation and decimation may be used to further downsample a sequence without incurring aliasing. It should be noted that maximum possible downsampling is achieved once the non-zero portion of one period of the discrete-time spectrum has expanded to fill the entire band from $-\pi$ to π.

Consider the sequence $x[n]$ whose Fourier transform $X(e^{j\omega})$ is illustrated in Figure 7.38(a). As discussed in Example 7.4, the lowest rate at which impulse-train sampling may be used on this sequence without incurring aliasing is $2\pi/4$. This corresponds to

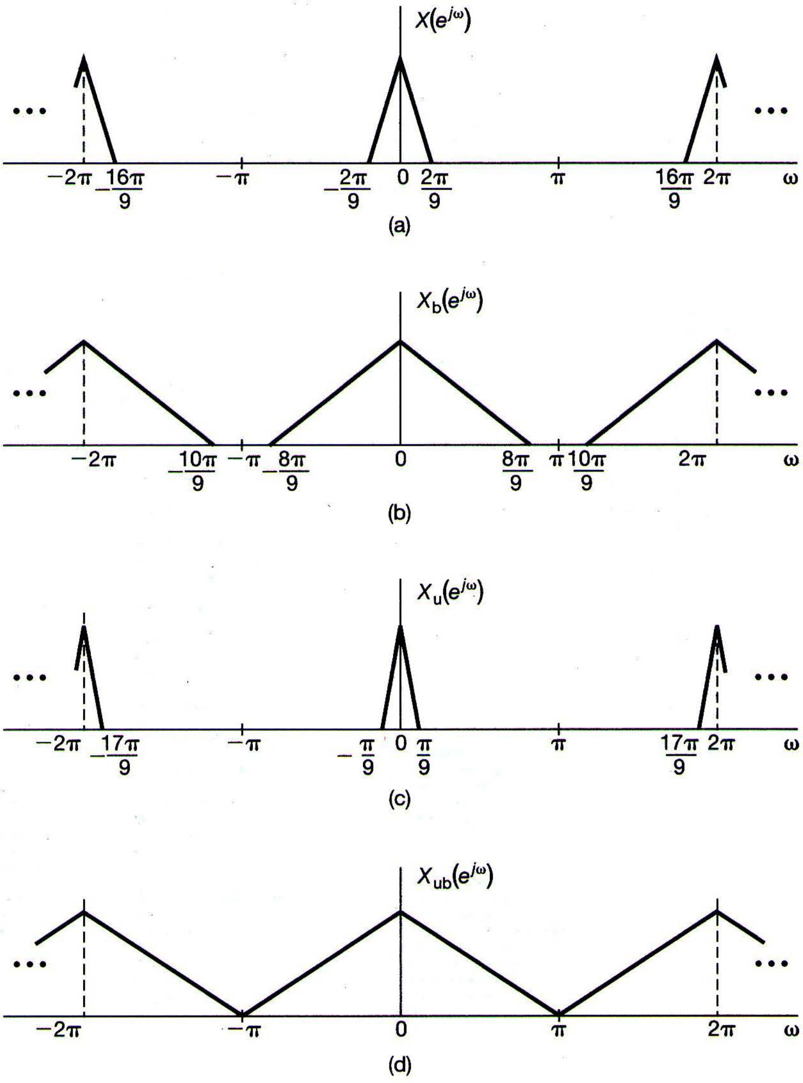


Figure 7.38 Spectra associated with Example 7.5. (a) Spectrum of $x[n]$; (b) spectrum after downsampling by 4; (c) spectrum after upsampling $x[n]$ by a factor of 2; (d) spectrum after upsampling $x[n]$ by 2 and then downsampling by 9.

sampling every 4th value of $x[n]$. If the result of such sampling is decimated by a factor of 4, we obtain a sequence $x_b[n]$ whose spectrum is shown in Figure 7.38(b). Clearly, there is still no aliasing of the original spectrum. However, this spectrum is zero for $8\pi/9 \leq |\omega| \leq \pi$, which suggests there is room for further downsampling.

Specifically, examining Figure 7.38(a) we see that if we could scale frequency by a factor of 9/2, the resulting spectrum would have nonzero values over the entire frequency interval from $-\pi$ to π. However, since 9/2 is not an integer, we can't achieve this purely by downsampling. Rather we must first upsample $x[n]$ by a factor of 2 and then downsample by a factor of 9. In particular, the spectrum of the signal $x_u[n]$ obtained when $x[n]$ is upsampled by a factor of 2, is displayed in Figure 7.38(c). When $x_u[n]$ is then downsampled by a factor of 9, the spectrum of the resulting sequence $x_{ub}[n]$ is as shown in Figure 7.38(d). This combined result effectively corresponds to downsampling $x[n]$ by a noninteger amount, 9/2. Assuming that $x[n]$ represents unaliased samples of a continuous-time signal $x_c(t)$, our interpolated and decimated sequence represents the maximum possible (aliasing-free) downsampling of $x_c(t)$.

7.6 SUMMARY

In this chapter we have developed the concept of sampling, whereby a continuous-time or discrete-time signal is represented by a sequence of equally spaced samples. The conditions under which the signal is exactly recoverable from the samples is embodied in the sampling theorem. For exact reconstruction, this theorem requires that the signal to be sampled be band limited and that the sampling frequency be greater than twice the highest frequency in the signal to be sampled. Under these conditions, exact reconstruction of the original signal is carried out by means of ideal lowpass filtering. The time-domain interpretation of this ideal reconstruction procedure is often referred to as ideal band-limited interpolation. In practical implementations, the lowpass filter is approximated and the interpolation in the time domain is no longer exact. In some instances, simple interpolation procedures such as a zero-order hold or linear interpolation (a first-order hold) suffice.

If a signal is undersampled (i.e., if the sampling frequency is less than that required by the sampling theorem), then the signal reconstructed by ideal band-limited interpolation will be related to the original signal through a form of distortion referred to as aliasing. In many instances, it is important to choose the sampling rate so as to avoid aliasing. However, there are a variety of important examples, such as the stroboscope, in which aliasing is exploited.

Sampling has a number of important applications. One particularly significant set of applications relates to using sampling to process continuous-time signals with discrete-time systems, by means of minicomputers, microprocessors, or any of a variety of devices specifically oriented toward discrete-time signal processing.

The basic theory of sampling is similar for both continuous-time and discrete-time signals. In the discrete-time case there is the closely related concept of decimation, whereby the decimated sequence is obtained by extracting values of the original sequence at equally spaced intervals. The difference between sampling and decimation lies in the fact that, for the sampled sequence, values of zero lie in between the sample values, whereas in the decimated sequence these zero values are discarded, thereby compressing the sequence in time. The inverse of decimation is interpolation. The ideas of decimation and interpolation arise in a variety of important practical applications of signals and systems, including communication systems, digital audio, high-definition television, and many other applications.

Chapter 7 Problems

The first section of problems belongs to the basic category, and the answers are provided in the back of the book. The remaining two sections contain problems belonging to the basic and advanced categories, respectively.

BASIC PROBLEMS WITH ANSWERS

7.1. A real-valued signal $x(t)$ is known to be uniquely determined by its samples when the sampling frequency is $\omega_s = 10{,}000\pi$. For what values of ω is $X(j\omega)$ guaranteed to be zero?

7.2. A continuous-time signal $x(t)$ is obtained at the output of an ideal lowpass filter with cutoff frequency $\omega_c = 1{,}000\pi$. If impulse-train sampling is performed on $x(t)$, which of the following sampling periods would guarantee that $x(t)$ can be recovered from its sampled version using an appropriate lowpass filter?

(a) $T = 0.5 \times 10^{-3}$

(b) $T = 2 \times 10^{-3}$

(c) $T = 10^{-4}$

7.3. The frequency which, under the sampling theorem, must be exceeded by the sampling frequency is called the *Nyquist rate*. Determine the Nyquist rate corresponding to each of the following signals:

(a) $x(t) = 1 + \cos(2{,}000\pi t) + \sin(4{,}000\pi t)$

(b) $x(t) = \frac{\sin(4{,}000\pi t)}{\pi t}$

(c) $x(t) = \left(\frac{\sin(4{,}000\pi t)}{\pi t}\right)^2$

7.4. Let $x(t)$ be a signal with Nyquist rate ω_0. Determine the Nyquist rate for each of the following signals:

(a) $x(t) + x(t-1)$

(b) $\frac{dx(t)}{dt}$

(c) $x^2(t)$

(d) $x(t)\cos\omega_0 t$

7.5. Let $x(t)$ be a signal with Nyquist rate ω_0. Also, let

$$y(t) = x(t)p(t-1),$$

where

$$p(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT), \text{ and } T < \frac{2\pi}{\omega_0}.$$

Specify the constraints on the magnitude and phase of the frequency response of a filter that gives $x(t)$ as its output when $y(t)$ is the input.

7.6. In the system shown in Figure P7.6, two functions of time, $x_1(t)$ and $x_2(t)$, are multiplied together, and the product $w(t)$ is sampled by a periodic impulse train. $x_1(t)$ is band limited to ω_1, and $x_2(t)$ is band limited to ω_2; that is,

$$X_1(j\omega) = 0, |\omega| \geq \omega_1,$$

$$X_2(j\omega) = 0, |\omega| \geq \omega_2.$$

Determine the *maximum* sampling interval T such that $w(t)$ is recoverable from $w_p(t)$ through the use of an ideal lowpass filter.

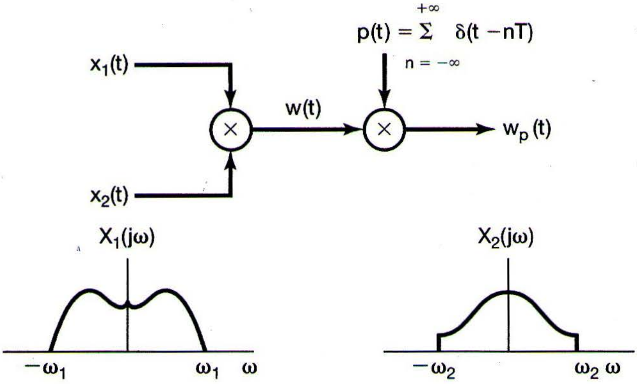


Figure P7.6

7.7. A signal $x(t)$ undergoes a zero-order hold operation with an effective sampling period T to produce a signal $x_0(t)$. Let $x_1(t)$ denote the result of a first-order hold operation on the samples of $x(t)$; i.e.,

$$x_1(t) = \sum_{n=-\infty}^{\infty} x(nT)h_1(t - nT),$$

where $h_1(t)$ is the function shown in Figure P7.7. Specify the frequency response of a filter that produces $x_1(t)$ as its output when $x_0(t)$ is the input.

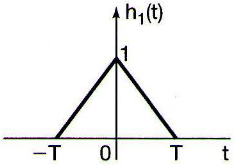


Figure P7.7

7.8. Consider a real, odd, and periodic signal $x(t)$ whose Fourier series representation may be expressed as

$$x(t) = \sum_{k=0}^{5} \left(\frac{1}{2}\right)^k \sin(k\pi t).$$

Let $\hat{x}(t)$ represent the signal obtained by performing impulse-train sampling on $x(t)$ using a sampling period of $T = 0.2$.

(a) Does aliasing occur when this impulse-train sampling is performed on $x(t)$?

(b) If $\hat{x}(t)$ is passed through an ideal lowpass filter with cutoff frequency π/T and passband gain T, determine the Fourier series representation of the output signal $g(t)$.

7.9. Consider the signal

$$x(t) = \left(\frac{\sin 50\pi t}{\pi t}\right)^2,$$

which we wish to sample with a sampling frequency of $\omega_s = 150\pi$ to obtain a signal $g(t)$ with Fourier transform $G(j\omega)$. Determine the maximum value of ω_0 for which it is guaranteed that

$$G(j\omega) = 75X(j\omega) \text{ for } |\omega| \leq \omega_0,$$

where $X(j\omega)$ is the Fourier transform of $x(t)$.

7.10. Determine whether each of the following statements is true or false:

(a) The signal $x(t) = u(t + T_0) - u(t - T_0)$ can undergo impulse-train sampling without aliasing, provided that the sampling period $T < 2T_0$.

(b) The signal $x(t)$ with Fourier transform $X(j\omega) = u(\omega + \omega_0) - u(\omega - \omega_0)$ can undergo impulse-train sampling without aliasing, provided that the sampling period $T < \pi/\omega_0$.

(c) The signal $x(t)$ with Fourier transform $X(j\omega) = u(\omega) - u(\omega - \omega_0)$ can undergo impulse-train sampling without aliasing, provided that the sampling period $T < 2\pi/\omega_0$.

7.11. Let $x_c(t)$ be a continuous-time signal whose Fourier transform has the property that $X_c(j\omega) = 0$ for $|\omega| \geq 2{,}000\pi$. A discrete-time signal

$$x_d[n] = x_c(n(0.5 \times 10^{-3}))$$

is obtained. For each of the following constraints on the Fourier transform $X_d(e^{j\omega})$ of $x_d[n]$, determine the corresponding constraint on $X_c(j\omega)$:

(a) $X_d(e^{j\omega})$ is real.

(b) The maximum value of $X_d(e^{j\omega})$ over all ω is 1.

(c) $X_d(e^{j\omega}) = 0$ for $\frac{3\pi}{4} \leq |\omega| \leq \pi$.

(d) $X_d(e^{j\omega}) = X_d(e^{j(\omega-\pi)})$.

7.12. A discrete-time signal $x_d[n]$ has a Fourier transform $X_d(e^{j\omega})$ with the property that $X_d(e^{j\omega}) = 0$ for $3\pi/4 \leq |\omega| \leq \pi$. The signal is converted into a continuous-time signal

$$x_c(t) = T \sum_{n=-\infty}^{\infty} x_d[n] \frac{\sin \frac{\pi}{T}(t - nT)}{\pi(t - nT)},$$

where $T = 10^{-3}$. Determine the values of ω for which the Fourier transform $X_c(j\omega)$ of $x_c(t)$ is guaranteed to be zero.

7.13. With reference to the filtering approach illustrated in Figure 7.24, assume that the sampling period used is T and the input $x_c(t)$ is band limited, so that $X_c(j\omega) = 0$ for $|\omega| \geq \pi/T$. If the overall system has the property that $y_c(t) = x_c(t-2T)$, determine the impulse response $h[n]$ of the discrete-time filter in Figure 7.24.

7.14. Repeat the previous problem, except this time assume that

$$y_c(t) = \frac{d}{dt} x_c\left(t - \frac{T}{2}\right).$$

7.15. Impulse-train sampling of $x[n]$ is used to obtain

$$g[n] = \sum_{k=-\infty}^{\infty} x[n]\delta[n - kN].$$

If $X(e^{j\omega}) = 0$ for $3\pi/7 \leq |\omega| \leq \pi$, determine the largest value for the sampling interval N which ensures that no aliasing takes place while sampling $x[n]$.

7.16. The following facts are given about the signal $x[n]$ and its Fourier transform:

1. $x[n]$ is real.
2. $X(e^{j\omega}) \neq 0$ for $0 < \omega < \pi$.
3. $x[n]\sum_{k=-\infty}^{\infty}\delta[n - 2k] = \delta[n]$.

Determine $x[n]$. You may find it useful to note that the signal $(\sin \frac{\pi}{2}n)/(\pi n)$ satisfies two of these conditions.

7.17. Consider an ideal discrete-time bandstop filter with impulse response $h[n]$ for which the frequency response in the interval $-\pi \le \omega \le \pi$ is

$$H(e^{j\omega}) = \begin{cases} 1, & |\omega| \le \frac{\pi}{4} \text{ and } |\omega| \ge \frac{3\pi}{4} \\ 0, & \text{elsewhere} \end{cases}.$$

Determine the frequency response of the filter whose impulse response is $h[2n]$.

7.18. Suppose the impulse response of an ideal discrete-time lowpass filter with cutoff frequency $\pi/2$ is interpolated (in accordance with Figure 7.37) to obtain an upsampling by a factor of 2. What is the frequency response corresponding to this upsampled impulse response?

7.19. Consider the system shown in Figure P7.19, with input $x[n]$ and the corresponding output $y[n]$. The zero-insertion system inserts two points with zero amplitude between each of the sequence values in $x[n]$. The decimation is defined by

$$y[n] = w[5n],$$

where $w[n]$ is the input sequence for the decimation system. If the input is of the form

$$x[n] = \frac{\sin \omega_1 n}{\pi n},$$

determine the output $y[n]$ for the following values of ω_1:

(a) $\omega_1 \le \frac{3\pi}{5}$

(b) $\omega_1 > \frac{3\pi}{5}$

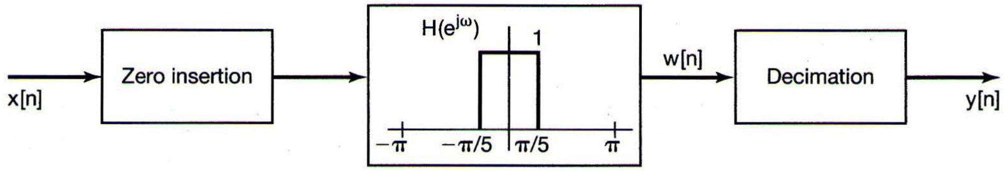


Figure P7.19

7.20. Two discrete-time systems S_1 and S_2 are proposed for implementing an ideal lowpass filter with cutoff frequency $\pi/4$. System S_1 is depicted in Figure P7.20(a). System S_2 is depicted in Figure P7.20(b). In these figures, S_A corresponds to a zero-insertion system that inserts one zero after every input sample, while S_B corresponds to a decimation system that extracts every second sample of its input.

(a) Does the proposed system S_1 correspond to the desired ideal lowpass filter?

(b) Does the proposed system S_2 correspond to the desired ideal lowpass filter?

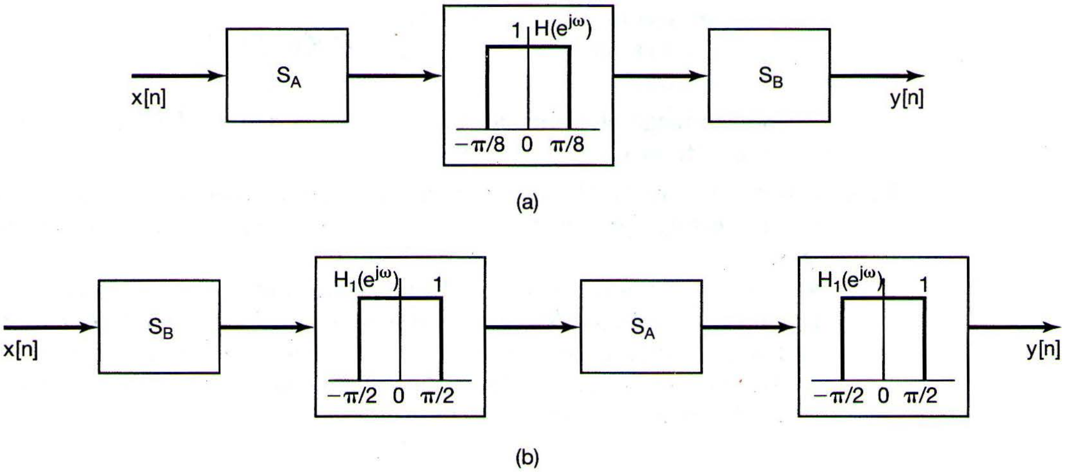


Figure P7.20

BASIC PROBLEMS

7.21. A signal $x(t)$ with Fourier transform $X(j\omega)$ undergoes impulse-train sampling to generate

$$x_p(t) = \sum_{n=-\infty}^{\infty} x(nT)\,\delta(t - nT)$$

where $T = 10^{-4}$. For each of the following sets of constraints on $x(t)$ and/or $X(j\omega)$, does the sampling theorem (see Section 7.1.1) guarantee that $x(t)$ can be recovered exactly from $x_p(t)$?

(a) $X(j\omega) = 0$ for $|\omega| > 5000\pi$
(b) $X(j\omega) = 0$ for $|\omega| > 15000\pi$
(c) $\mathcal{R}e\{X(j\omega)\} = 0$ for $|\omega| > 5000\pi$
(d) $x(t)$ real and $X(j\omega) = 0$ for $\omega > 5000\pi$
(e) $x(t)$ real and $X(j\omega) = 0$ for $\omega < -15000\pi$
(f) $X(j\omega) * X(j\omega) = 0$ for $|\omega| > 15000\pi$
(g) $|X(j\omega)| = 0$ for $\omega > 5000\pi$

7.22. The signal $y(t)$ is generated by convolving a band-limited signal $x_1(t)$ with another band-limited signal $x_2(t)$, that is,

$$y(t) = x_1(t) * x_2(t)$$

where

$$X_1(j\omega) = 0 \qquad \text{for } |\omega| > 1000\pi$$
$$X_2(j\omega) = 0 \qquad \text{for } |\omega| > 2000\pi.$$

Impulse-train sampling is performed on $y(t)$ to obtain

$$y_p(t) = \sum_{n=-\infty}^{+\infty} y(nT)\delta(t - nT).$$

Specify the range of values for the sampling period T which ensures that $y(t)$ is recoverable from $y_p(t)$.

7.23. Shown in Figure P7.23 is a system in which the sampling signal is an impulse train with alternating sign. The Fourier transform of the input signal is as indicated in the figure.

(a) For $\Delta < \pi/(2\omega_M)$, sketch the Fourier transform of $x_p(t)$ and $y(t)$.
(b) For $\Delta < \pi/(2\omega_M)$, determine a system that will recover $x(t)$ from $x_p(t)$.
(c) For $\Delta < \pi/(2\omega_M)$, determine a system that will recover $x(t)$ from $y(t)$.
(d) What is the *maximum* value of Δ in relation to ω_M for which $x(t)$ can be recovered from either $x_p(t)$ or $y(t)$?

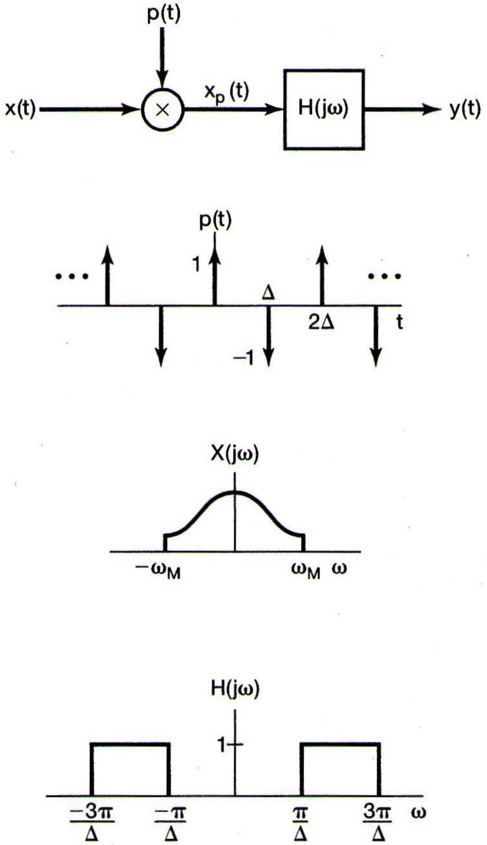


Figure P7.23

7.24. Shown in Figure P7.24 is a system in which the input signal is multiplied by a periodic square wave. The period of $s(t)$ is T. The input signal is band limited with $|X(j\omega)| = 0$ for $|\omega| \geq \omega_M$.

(a) For $\Delta = T/3$, determine, in terms of ω_M, the maximum value of T for which there is no aliasing among the replicas of $X(j\omega)$ in $W(j\omega)$.

(b) For $\Delta = T/4$, determine, in terms of ω_M, the maximum value of T for which there is no aliasing among the replicas of $X(j\omega)$ in $W(j\omega)$.

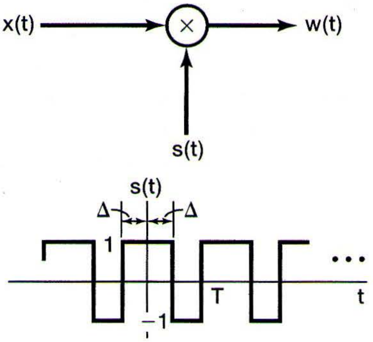


Figure P7.24

7.25. In Figure P7.25 is a sampler, followed by an ideal lowpass filter, for reconstruction of $x(t)$ from its samples $x_p(t)$. From the sampling theorem, we know that if $\omega_s = 2\pi/T$ is greater than twice the highest frequency present in $x(t)$ and $\omega_c = \omega_s/2$, then the reconstructed signal $x_r(t)$ will exactly equal $x(t)$. If this condition on the bandwidth of $x(t)$ is violated, then $x_r(t)$ will *not* equal $x(t)$. We seek to show in this problem that if $\omega_c = \omega_s/2$, then for any choice of T, $x_r(t)$ and $x(t)$ will always be equal at the sampling instants; that is,

$$x_r(kT) = x(kT),\ k = 0, \pm 1, \pm 2, \ldots.$$

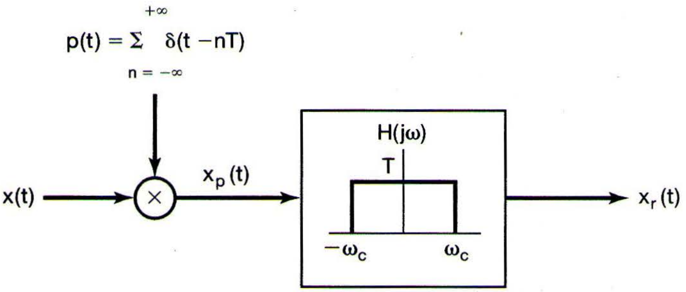


Figure P7.25

To obtain this result, consider eq. (7.11), which expresses $x_r(t)$ in terms of the samples of $x(t)$:

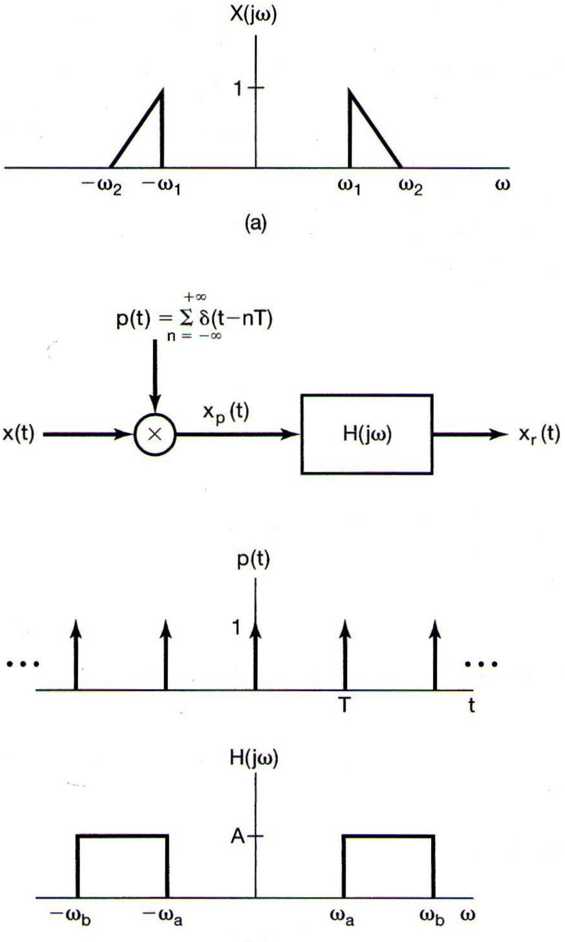
$$x_r(t) = \sum_{n=-\infty}^{\infty} x(nT) T \frac{\omega_c}{\pi} \frac{\sin[\omega_c(t - nT)]}{\omega_c(t - nT)}.$$

With $\omega_c = \omega_s/2$, this becomes

$$x_r(t) = \sum_{n=-\infty}^{\infty} x(nT) \frac{\sin\left[\frac{\pi}{T}(t - nT)\right]}{\frac{\pi}{T}(t - nT)}. \qquad \text{(P7.25–1)}$$

By considering the values of α for which $[\sin(\alpha)]/\alpha = 0$, show from eq. (P7.25–1) that, without any restrictions on $x(t)$, $x_r(kT) = x(kT)$ for any integer value of k.

7.26. The sampling theorem, as we have derived it, states that a signal $x(t)$ must be sampled at a rate greater than its bandwidth (or equivalently, a rate greater than twice its highest frequency). This implies that if $x(t)$ has a spectrum as indicated in Figure P7.26(a) then $x(t)$ must be sampled at a rate greater than $2\omega_2$. However, since the signal has most of its energy concentrated in a narrow band, it would seem reasonable to expect that a sampling rate lower than twice the highest frequency could be used. A signal whose energy is concentrated in a frequency band is often referred to as a *bandpass signal*. There are a variety of techniques for sampling such signals, generally referred to as *bandpass-sampling* techniques.



(a)

(b)

Figure P7.26

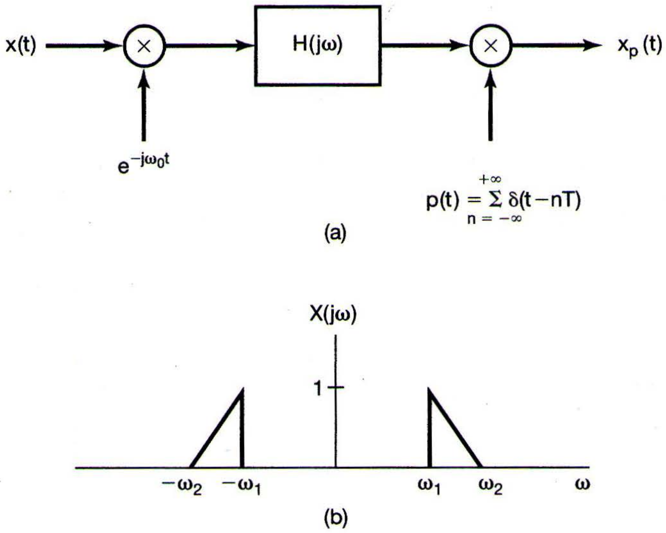
To examine the possibility of sampling a bandpass signal as a rate less than the total bandwidth, consider the system shown in Figure P7.26(b). Assuming that $\omega_1 > \omega_2 - \omega_1$, find the maximum value of T and the values of the constants A, ω_a, and ω_b such that $x_r(t) = x(t)$.

7.27. In Problem 7.26, we considered one procedure for bandpass sampling and reconstruction. Another procedure, used when $x(t)$ is real, consists of multiplying $x(t)$ by a complex-exponential and then sampling the product. The sampling system is shown in Figure P7.27(a). With $x(t)$ real and with $X(j\omega)$ nonzero only for $\omega_1 < |\omega| < \omega_2$, the frequency is chosen to be $\omega_0 = (1/2)(\omega_1 + \omega_2)$, and the lowpass filter $H_1(j\omega)$ has cutoff frequency $(1/2)(\omega_2 - \omega_1)$.

(a) For $X(j\omega)$ as shown in Figure P7.27(b), sketch $X_p(j\omega)$.

(b) Determine the maximum sampling period T such that $x(t)$ is recoverable from $x_p(t)$.

(c) Determine a system to recover $x(t)$ from $x_p(t)$.



(a)

(b)

Figure P7.27

7.28. Figure P7.28(a) shows a system that converts a continuous-time signal to a discrete-time signal. The input $x(t)$ is periodic with a period of 0.1 second. The Fourier series coefficients of $x(t)$ are

$$a_k = \left(\frac{1}{2}\right)^{|k|}, \quad -\infty < k < +\infty.$$

The lowpass filter $H(j\omega)$ has the frequency response shown in Figure P7.28(b). The sampling period $T = 5 \times 10^{-3}$ second.

(a) Show that $x[n]$ is a periodic sequence, and determine its period.

(b) Determine the Fourier series coefficients of $x[n]$.

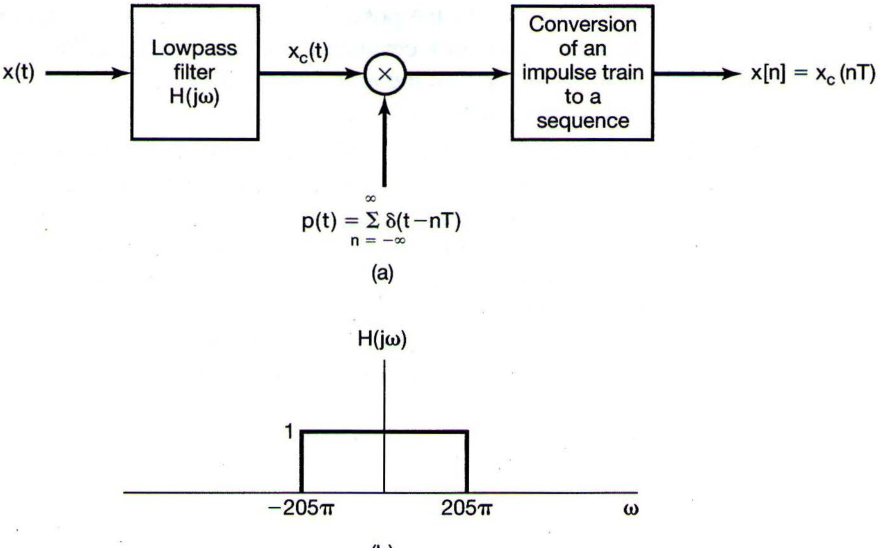

x(t)
Lowpass filter H(jω)
x_c(t)
×
Conversion of an impulse train to a sequence
x[n] = x_c(nT)
p(t) = Σ δ(t−nT), n = −∞ to ∞
(a)
H(jω)
1
−205π
205π
ω
(b)

Figure P7.28

7.29. Figure P7.29(a) shows the overall system for filtering a continuous-time signal using a discrete-time filter. If $X_c(j\omega)$ and $H(e^{j\omega})$ are as shown in Figure P7.29(b), with $1/T = 20$ kHz, sketch $X_p(j\omega)$, $X(e^{j\omega})$, $Y(e^{j\omega})$, $Y_p(j\omega)$, and $Y_c(j\omega)$.

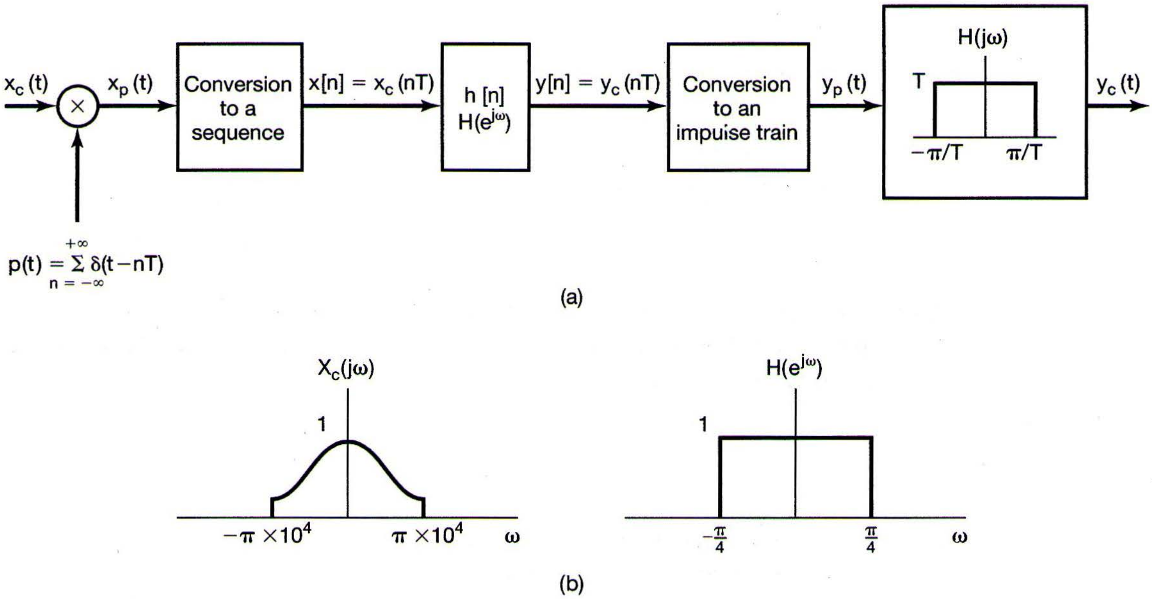

x_c(t)
×
x_p(t)
Conversion to a sequence
x[n] = x_c(nT)
h[n]
H(e^{jω})
y[n] = y_c(nT)
Conversion to an impulse train
y_p(t)
H(jω)
T
−π/T
π/T
y_c(t)
p(t) = Σ δ(t−nT), n = −∞ to +∞
(a)
X_c(jω)
1
−π ×10^4
π ×10^4
ω
H(e^{jω})
1
−π/4
π/4
ω
(b)

Figure P7.29

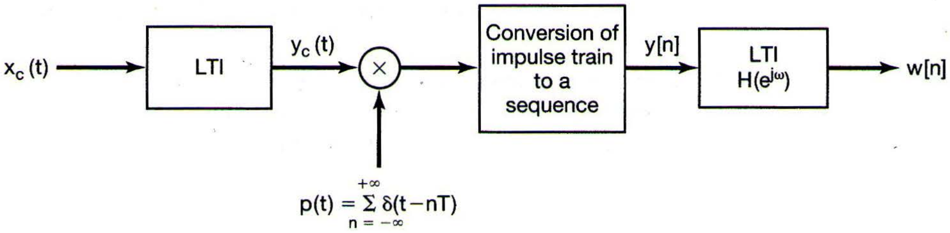
7.30. Figure P7.30 shows a system consisting of a continuous-time LTI system followed by a sampler, conversion to a sequence, and an LTI discrete-time system. The continuous-time LTI system is causal and satisfies the linear, constant-coefficient differential equation

$$\frac{dy_c(t)}{dt} + y_c(t) = x_c(t).$$

The input $x_c(t)$ is a unit impulse $\delta(t)$.

(a) Determine $y_c(t)$.

(b) Determine the frequency response $H(e^{j\omega})$ and the impulse response $h[n]$ such that $w[n] = \delta[n]$.



x_c (t)
LTI
y_c (t)
Conversion of impulse train to a sequence
y[n]
LTI H(e^{jω})
w[n]
p(t) = Σ_{n=-∞}^{+∞} δ(t−nT)

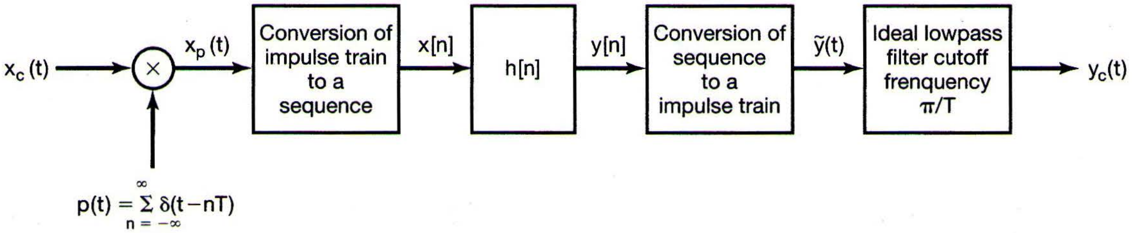
Figure P7.30

7.31. Shown in Figure P7.31 is a system that processes continuous-time signals using a digital filter $h[n]$ that is linear and causal with difference equation

$$y[n] = \frac{1}{2}y[n-1] + x[n].$$

For input signals that are band limited such that $X_c(j\omega) = 0$ for $|\omega| > \pi/T$, the system in the figure is equivalent to a continuous-time LTI system.

Determine the frequency response $H_c(j\omega)$ of the equivalent overall system with input $x_c(t)$ and output $y_c(t)$.



x_c (t)
x_p (t)
Conversion of impulse train to a sequence
x[n]
h[n]
y[n]
Conversion of sequence to a impulse train
ỹ(t)
Ideal lowpass filter cutoff frenquency π/T
y_c(t)
p(t) = Σ_{n=-∞}^{∞} δ(t−nT)

Figure P7.31

7.32. A signal $x[n]$ has a Fourier transform $X(e^{j\omega})$ that is zero for $(\pi/4) \le |\omega| \le \pi$. Another signal

$$g[n] = x[n] \sum_{k=-\infty}^{\infty} \delta[n-1-4k]$$

is generated. Specify the frequency response $H(e^{j\omega})$ of a lowpass filter that produces $x[n]$ as output when $g[n]$ is the input.

7.33. A signal $x[n]$ with Fourier transform $X(e^{j\omega})$ has the property that

$$\left(x[n] \sum_{k=-\infty}^{\infty} \delta[n-3k]\right) * \left(\frac{\sin \frac{\pi}{3} n}{\frac{\pi}{3} n}\right) = x[n].$$

For what values of ω is it guaranteed that $X(e^{j\omega}) = 0$?

7.34. A real-valued discrete-time signal $x[n]$ has a Fourier transform $X(e^{j\omega})$ that is zero for $3\pi/14 \le |\omega| \le \pi$. The nonzero portion of the Fourier transform of one period of $X(e^{j\omega})$ can be made to occupy the region $|\omega| < \pi$ by first performing upsampling by a factor of L and then performing downsampling by a factor of M. Specify the values of L and M.

7.35. Consider a discrete-time sequence $x[n]$ from which we form two new sequences, $x_p[n]$ and $x_d[n]$, where $x_p[n]$ corresponds to sampling $x[n]$ with a sampling period of 2 and $x_d[n]$ corresponds to decimating $x[n]$ by a factor of 2, so that

$$x_p[n] = \begin{cases} x[n], & n = 0, \pm 2, \pm 4, \ldots \\ 0, & n = \pm 1, \pm 3, \ldots \end{cases}$$

and

$$x_d[n] = x[2n].$$

(a) If $x[n]$ is as illustrated in Figure P7.35(a), sketch the sequences $x_p[n]$ and $x_d[n]$.
(b) If $X(e^{j\omega})$ is as shown in Figure P7.35(b), sketch $X_p(e^{j\omega})$ and $X_d(e^{j\omega})$.

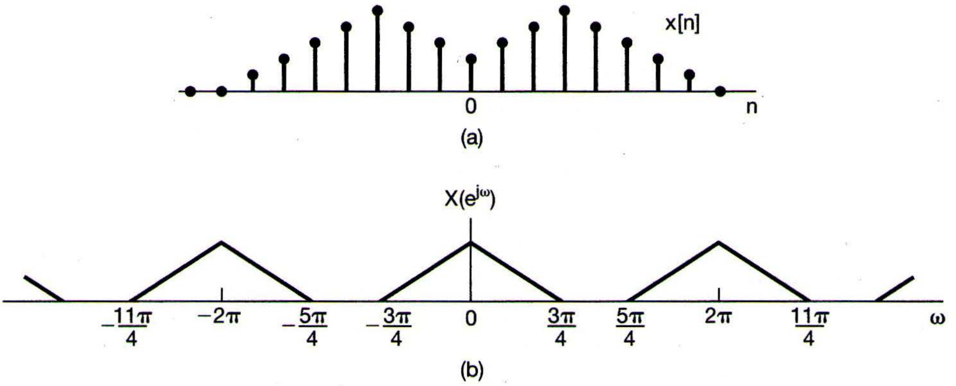


Figure P7.35

ADVANCED PROBLEMS

7.36 Let $x(t)$ be a band-limited signal such that $X(j\omega) = 0$ for $|\omega| \ge \frac{\pi}{T}$.
(a) If $x(t)$ is sampled using a sampling period T, determine an interpolating function

$g(t)$ such that

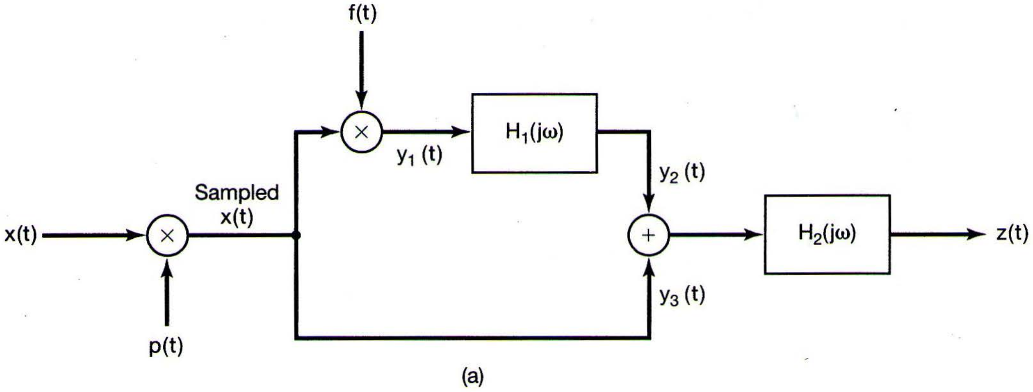
$$\frac{dx(t)}{dt} = \sum_{n=-\infty}^{\infty} x(nT)g(t - nT).$$

(b) Is the function $g(t)$ unique?

7.37. A signal limited in bandwidth to $|\omega| < W$ can be recovered from nonuniformly spaced samples as long as the average sample density is $2(W/2\pi)$ samples per second. This problem illustrates a particular example of nonuniform sampling. Assume that in Figure P7.37(a):

1. $x(t)$ is band limited; $X(j\omega) = 0, |\omega| > W$.
2. $p(t)$ is a nonuniformly spaced periodic pulse train, as shown in Figure P7.37(b).
3. $f(t)$ is a periodic waveform with period $T = 2\pi/W$. Since $f(t)$ multiplies an impulse train, only its values $f(0) = a$ and $f(\Delta) = b$ at $t = 0$ and $t = \Delta$, respectively, are significant.
4. $H_1(j\omega)$ is a 90° phase shifter; that is,

$$H_1(j\omega) = \begin{cases} j, & \omega > 0 \\ -j, & \omega < 0 \end{cases}.$$


f(t)
×
y_1 (t)
$H_1(j\omega)$
y_2 (t)
Sampled x(t)
x(t)
×
+
$H_2(j\omega)$
z(t)
y_3 (t)
p(t)
(a)

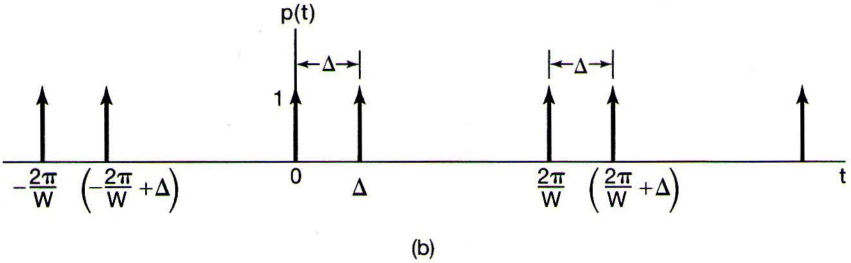

p(t)
Δ
Δ
1
$-\frac{2\pi}{W}$
$\left(-\frac{2\pi}{W}+\Delta\right)$
0
Δ
$\frac{2\pi}{W}$
$\left(\frac{2\pi}{W}+\Delta\right)$
t
(b)

Figure P7.37

5. $H_2(j\omega)$ is an ideal lowpass filter; that is,

$$H_2(j\omega) = \begin{cases} K, & 0 < \omega < W \\ K^*, & -W < \omega < 0\,, \\ 0, & |\omega| > W \end{cases}$$

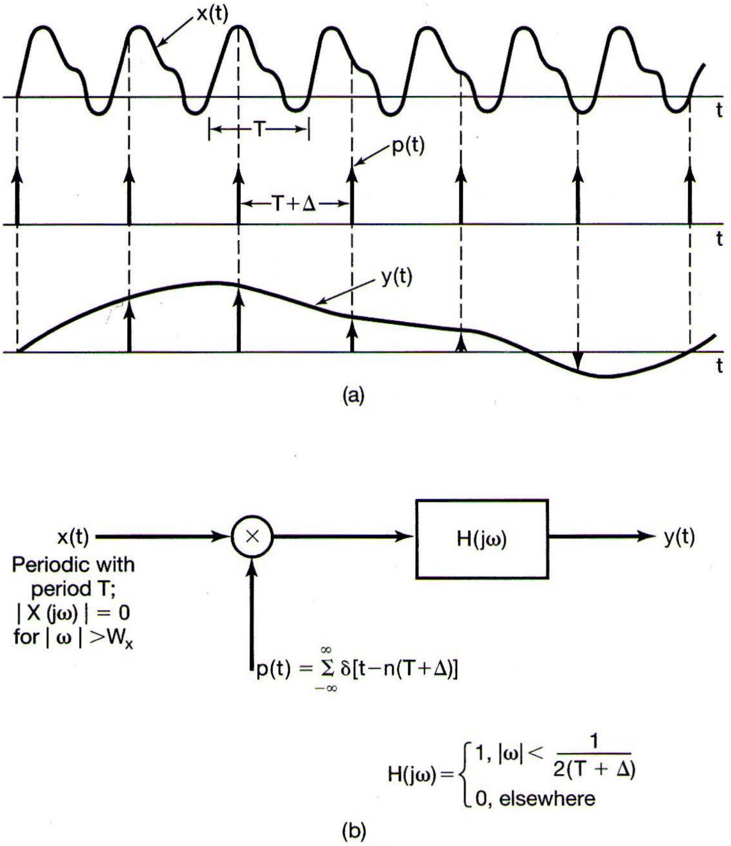
where K is a (possibly complex) constant.

(a) Find the Fourier transforms of $p(t)$, $y_1(t)$, $y_2(t)$, and $y_3(t)$.

(b) Specify the values of a, b, and K as functions of Δ such that $z(t) = x(t)$ for any band-limited $x(t)$ and any Δ such that $0 < \Delta < \pi/W$.

7.38. It is frequently necessary to display on an oscilloscope screen waveforms having very short time structures—for example, on the scale of thousandths of a nanosecond. Since the rise time of the fastest oscilloscope is longer than this, such displays cannot be achieved directly. If however, the waveform is periodic, the desired result can be obtained indirectly by using an instrument called a sampling oscilloscope.

The idea, as shown in Figure P7.38(a), is to sample the fast waveform $x(t)$ once each period, but at successively later points in successive periods. The increment Δ should be an appropriately chosen sampling interval in relation to the bandwidth of $x(t)$. If the resulting impulse train is then passed through an appropriate interpolat-



(a)

(b)

Figure P7.38

ing lowpass filter, the output $y(t)$ will be proportional to the original fast waveform slowed down or stretched out in time [i.e., $y(t)$ is proportional to $x(at)$, where $a < 1$].

For $x(t) = A + B\cos[(2\pi/T)t + \theta]$, find a range of values of Δ such that $y(t)$ in Figure P7.38(b) is proportional to $x(at)$ with $a < 1$. Also, determine the value of a in terms of T and Δ.

7.39 A signal $x_p(t)$ is obtained through impule-train sampling of a sinusoidal signal $x(t)$ whose frequency is equal to half the sampling frequency ω_s.

$$x(t) = \cos\left(\frac{\omega_s}{2}t + \phi\right)$$

and

$$x_p(t) = \sum_{n=-\infty}^{+\infty} x(nT)\delta(t - nT)$$

where $T = 2\pi/\omega_s$.

(a) Find $g(t)$ such that

$$x(t) = \cos(\phi)\cos\left(\frac{\omega_s}{2}t\right) + g(t).$$

(b) Show that

$$g(nT) = 0 \qquad \text{for } n = 0, \pm1, \pm2, \cdots$$

(c) Using the results of the previous two parts, show that if $x_p(t)$ is applied as the input to an ideal lowpass filter with cutoff frequency $\omega_s/2$, the resulting output is

$$y(t) = \cos(\phi)\cos\left(\frac{\omega_s}{2}t\right).$$

7.40. Consider a disc on which four cycles of a sinusoid are painted. The disc is rotated at approximately 15 revolutions per second, so that the sinusoid, when viewed through a narrow slit, has a frequency of 60 Hz.

The arrangement is indicated in Figure P7.40. Let $v(t)$ denote the position of the line seen through the slit. Then

$$v(t) = A\cos(\omega_0 t + \phi),\ \omega_0 = 120\pi.$$

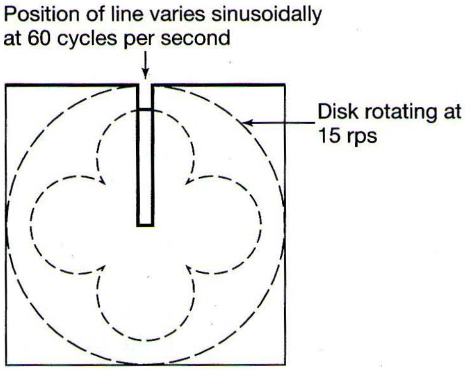

Position of line varies sinusoidally at 60 cycles per second
Disk rotating at 15 rps

Figure P7.40

For notational convenience, we will normalize $v(t)$ so that $A = 1$. At 60 Hz, the eye is not able to follow $v(t)$, and we will assume that this effect can be explained by modeling the eye as an ideal lowpass filter with cutoff frequency 20 Hz.

Sampling of the sinusoid can be accomplished by illuminating the disc with a strobe light. Thus, the illumination can be represented by an impulse train; that is,

$$i(t) = \sum_{k=-\infty}^{+\infty} \delta(t - kT),$$

where $1/T$ is the strobe frequency in hertz. The resulting sampled signal is the product $r(t) = v(t)i(t)$. Let $R(j\omega)$, $V(j\omega)$, and $I(j\omega)$ denote the Fourier transforms of $r(t)$, $v(t)$, and $i(t)$, respectively.

(a) Sketch $V(j\omega)$, indicating clearly the effect of the parameters ϕ and ω_0.

(b) Sketch $I(j\omega)$, indicating the effect of T.

(c) According to the sampling theorem, there is a maximum value for T in terms of ω_0 such that $v(t)$ can be recovered from $r(t)$ using a lowpass filter. Determine this value of T and the cutoff frequency of the lowpass filter. Sketch $R(j\omega)$ when T is slightly less than the maximum value.

If the sampling period T is made greater than the value determined in part (c), aliasing of the spectrum occurs. As a result of this aliasing, we perceive a lower frequency sinusoid.

(d) Suppose that $2\pi/T = \omega_0 + 20\pi$. Sketch $R(j\omega)$ for $|\omega| < 40\pi$. Denote by $v_a(t)$ the apparent position of the line as we perceive it. Assuming that the eye behaves as an ideal lowpass filter with 20-Hz cutoff and unity gain, express $v_a(t)$ in the form

$$v_a(t) = A_a \cos(\omega_a + \phi_a),$$

where A_a is the apparent amplitude, ω_a the apparent frequency, and ϕ_a the apparent phase of $v_a(t)$.

(e) Repeat part (d) for $2\pi/T = \omega_0 - 20\pi$.

7.41. In many practical situations a signal is recorded in the presence of an echo, which we would like to remove by appropriate processing. For example, in Figure P7.41(a), we illustrate a system in which a receiver simultaneously receives a signal $x(t)$ and an echo represented by an attenuated delayed replication of $x(t)$. Thus, the receiver output is $s(t) = x(t) + \alpha x(t - T_0)$, where $|\alpha| < 1$. This output is to be processed to recover $x(t)$ by first converting to a sequence and then using an appropriate digital filter $h[n]$, as indicated in Figure P7.41(b).

Assume that $x(t)$ is band limited [i.e., $X(j\omega) = 0$ for $|\omega| > \omega_M$] and that $|\alpha| < 1$.

(a) If $T_0 < \pi/\omega_M$, and the sampling period is taken to be equal to T_0 (i.e., $T = T_0$), determine the difference equation for the digital filter $h[n]$ such that $y_c(t)$ is proportional to $x(t)$.

(b) With the assumptions of part (a), specify the gain A of the ideal lowpass filter such that $y_c(t) = x(t)$.

(c) Now suppose that $\pi/\omega_M < T_0 < 2\pi/\omega_M$. Determine a choice for the sampling period T, the lowpass filter gain A, and the frequency response for the digital filter $h[n]$ such that $y_c(t)$ is proportional to $x(t)$.

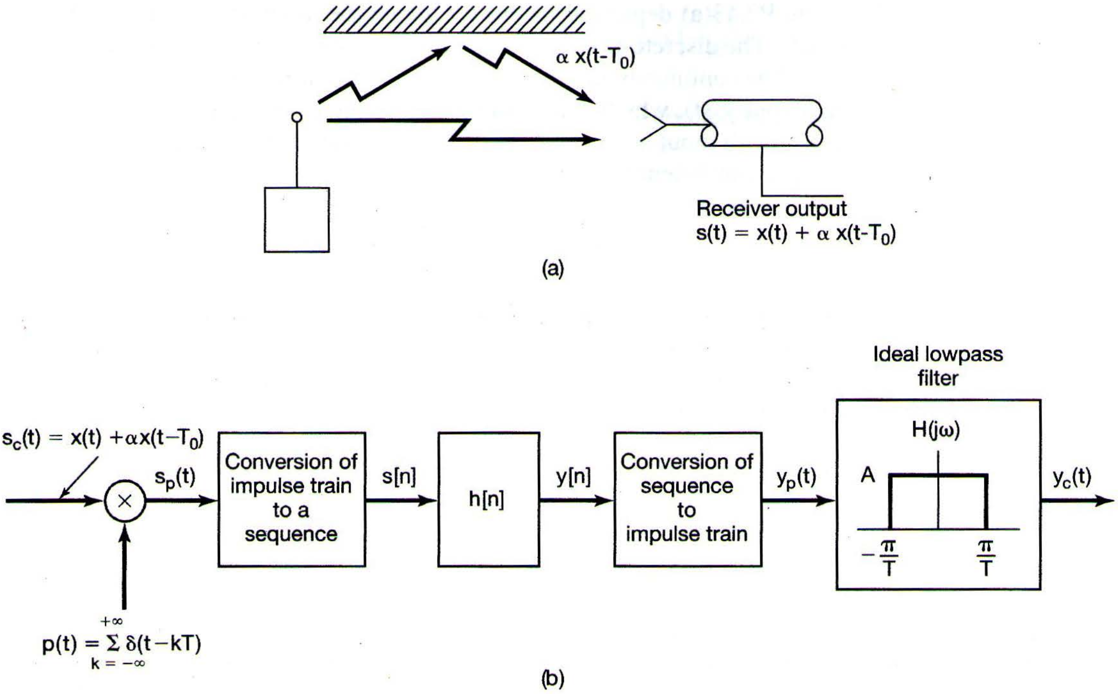

α x(t-T_0)
Receiver output
s(t) = x(t) + α x(t-T_0)
(a)
Ideal lowpass filter
s_c(t) = x(t) + αx(t−T_0)
s_p(t)
Conversion of impulse train to a sequence
s[n]
h[n]
y[n]
Conversion of sequence to impulse train
y_p(t)
H(jω)
A
−π/T
π/T
y_c(t)
p(t) = Σ_{k=−∞}^{+∞} δ(t−kT)
(b)

Figure P7.41

7.42. Consider a band-limited signal $x_c(t)$ that is sampled at a rate higher than the Nyquist rate. The samples, spaced T seconds apart, are then converted to a sequence $x[n]$, as indicated in Figure P7.42.

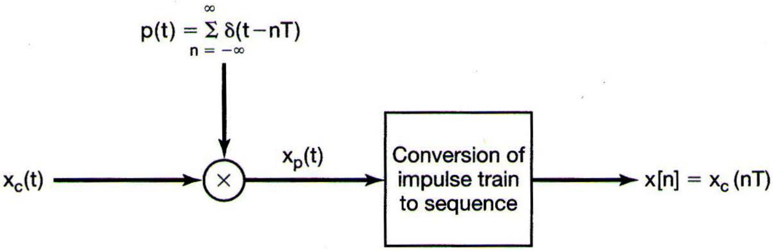

p(t) = Σ_{n=−∞}^{∞} δ(t−nT)
x_c(t)
x_p(t)
Conversion of impulse train to sequence
x[n] = x_c (nT)

Figure P7.42

Determine the relation between the energy E_d of the sequence, the energy E_c of the original signal, and the sampling interval T. The energy of a sequence $x[n]$ is defined as

$$E_d = \sum_{n=-\infty}^{\infty} |x[n]|^2,$$

and the energy in a continuous-time function $x_c(t)$ is defined as

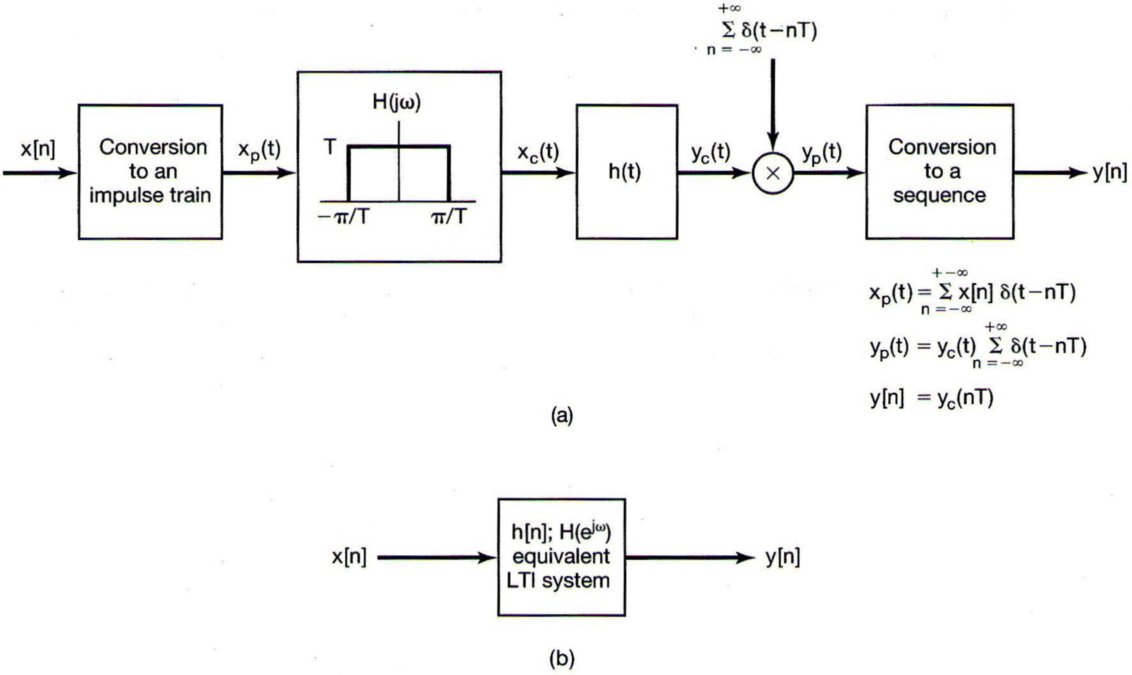
$$E_c = \int_{-\infty}^{+\infty} |x_c(t)|^2\, dt.$$

7.43. Figure P7.43(a) depicts a system for which the input and output are discrete-time signals. The discrete-time input $x[n]$ is converted to a continuous-time impulse train $x_p(t)$. The continuous-time signal $x_p(t)$ is then filtered by an LTI system to produce the output $y_c(t)$, which is then converted to the discrete-time signal $y[n]$. The LTI system with input $x_c(t)$ and output $y_c(t)$ is causal and is characterized by the linear constant-coefficient differential equation

$$\frac{d^2y_c(t)}{dt^2} + 4\frac{dy_c(t)}{dt} + 3y_c(t) = x_c(t).$$

The overall system is equivalent to a causal discrete-time LTI system, as indicated in Figure P7.43(b).

Determine the frequency response $H(e^{j\omega})$ and the unit sample response $h[n]$ of the equivalent LTI system.



x[n]
Conversion to an impulse train
x_p(t)
H(jω)
T
−π/T
π/T
x_c(t)
h(t)
y_c(t)
Σ δ(t−nT), n = −∞ to +∞
×
y_p(t)
Conversion to a sequence
y[n]
x_p(t) = Σ x[n] δ(t−nT), n = −∞ to +∞
y_p(t) = y_c(t) Σ δ(t−nT), n = −∞ to +∞
y[n] = y_c(nT)
x[n]
h[n]; H(e^{jω}) equivalent LTI system
y[n]

(a)

(b)

Figure P7.43

7.44. Suppose we wish to design a continuous-time generator that is capable of producing sinusoidal signals at any frequency satisfying

$$\omega_1 \leq \omega \leq \omega_2,$$

where ω_1 and ω_2 are given positive numbers.

Our design is to take the following form: We have stored a discrete-time cosine wave of period N; that is, we have stored $x[0], \ldots, x[N-1]$, where

$$x[k] = \cos\left(\frac{2\pi k}{N}\right).$$

Every T seconds we output an impulse weighted by a value of $x[k]$, where we proceed through the values of $k = 0, 1, \ldots, N-1$ in a cyclic fashion. That is,

$$y_p(kT) = x(k \text{ modulo } N),$$

or equivalently,

$$y_p(kT) = \cos\left(\frac{2\pi k}{N}\right),$$

and

$$y_p(t) = \sum_{k=-\infty}^{+\infty} \cos\left(\frac{2\pi k}{N}\right) \delta(t - kT).$$

(a) Show that by adjusting T, we can adjust the frequency of the cosine signal being sampled. That is, show that

$$y_p(t) = (\cos \omega_0 t) \sum_{k=-\infty}^{+\infty} \delta(t - kT),$$

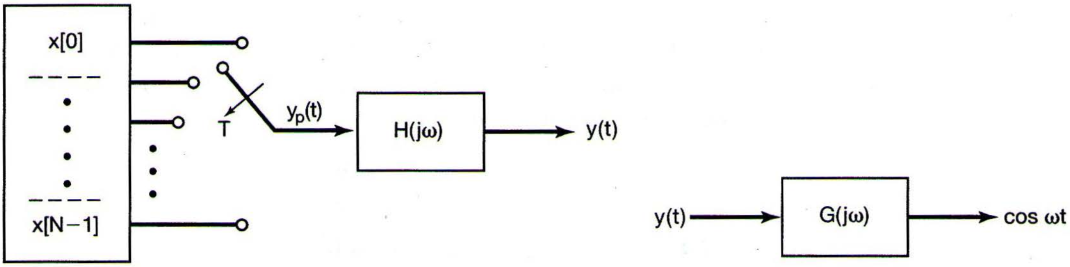
where $\omega_0 = 2\pi/NT$. Determine a range of values for T such that $y_p(t)$ can represent samples of a cosine signal with a frequency that is variable over the full range

$$\omega_1 \leq \omega \leq \omega_2.$$

(b) Sketch $Y_p(j\omega)$.

The overall system for generating a continuous-time sinusoid is depicted in Figure P7.44(a). $H(j\omega)$ is an ideal lowpass filter with unity gain in its passband; that is,

$$H(j\omega) = \begin{cases} 1, & |\omega| < \omega_c \\ 0, & \text{otherwise} \end{cases}.$$



(a) (b)

Figure P7.44

The parameter ω_c is to be determined so that $y(t)$ is a continuous-time cosine signal in the desired frequency band.

(c) Consider any value of T in the range determined in part (a). Determine the minimum value of N and some value for ω_c such that $y(t)$ is a cosine signal in the range $\omega_1 \leq \omega \leq \omega_2$.

(d) The amplitude of $y(t)$ will vary, depending upon the value of ω chosen between ω_1 and ω_2. Thus, we must design a system $G(j\omega)$ that normalizes the signal as shown in Figure P7.44(b). Find such a $G(j\omega)$.

7.45. In the system shown in Figure P7.45, the input $x_c(t)$ is band limited with $X_c(j\omega) = 0, |\omega| > 2\pi \times 10^4$. The digital filter $h[n]$ is described by the input-output relation

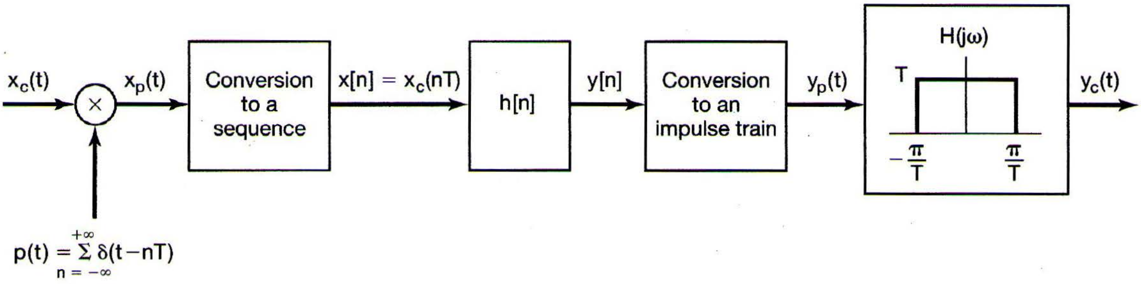


Figure P7.45

$$y[n] = T \sum_{k=-\infty}^{n} x[k]. \tag{P7.45–1}$$

(a) What is the maximum value of T allowed if aliasing is to be avoided in the transformation from $x_c(t)$ to $x_p(t)$.

(b) With the discrete-time LTI system specified through eq. (P7.45–1), determine its impulse response $h[n]$.

(c) Determine whether there is any value of T for which

$$\lim_{n\to\infty} y[n] = \lim_{t\to\infty} \int_{-\infty}^{t} x_c(\tau)\,d\tau. \tag{P7.45–2}$$

If so, determine the *maximum* value. If not, explain and specify how T would be chosen so that the equality in eq. (P7.45–2) is best approximated. (Think carefully about this part; it is easy to jump to the wrong conclusion!)

7.46 A signal $x[n]$ is sampled in discrete time as shown in Figure P7.46. $h_r[n]$ is an ideal lowpass filter with frequency response

$$H_r(e^{j\omega}) = \begin{cases} 1, & |\omega| < \frac{\pi}{N} \\ 0, & \frac{\pi}{N} < |\omega| < \pi \end{cases}$$

From eqs. (7.46) and (7.47), the filter output is expressible as

$$x_r[n] = \sum_{k=-\infty}^{+\infty} x[kN]h_r[n-kN] = \sum_{k=-\infty}^{+\infty} x[kN]\frac{N\omega_c}{\pi}\frac{\sin\omega_c(n-kN)}{\omega_c(n-kN)}$$

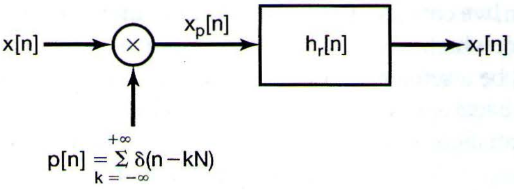


Figure P7.46

where $\omega_c = 2\pi/N$. Show that independent of whether the sequence $x[n]$ is sampled above or below the Nyquist rate, $x_r[mN] = x[mN]$, where m is any positive or negative integer.

7.47. Suppose $x[n]$ has a Fourier transform that is zero for $\pi/3 \leq |\omega| \leq \pi$. Show that

$$x[n] = \sum_{k=-\infty}^{\infty} x[3k]\left(\frac{\sin(\frac{\pi}{3}(n-3k))}{\frac{\pi}{3}(n-3k)}\right).$$

7.48. If $x[n] = \cos(\frac{\pi}{4}n+\phi_0)$ with $0 \leq \phi_0 < 2\pi$ and $g[n] = x[n]\sum_{k=-\infty}^{\infty}\delta[n-4k]$, what additional constraints must be imposed on ϕ_0 to ensure that

$$g[n] * \left(\frac{\sin\frac{\pi}{4}n}{\frac{\pi}{4}n}\right) = x[n]?$$

7.49. As discussed in Section 7.5 and illustrated in Figure 7.37, the procedure for interpolation or upsampling by an integer factor N can be thought of as the cascade of two operations. The first operation, involving system A, corresponds to inserting $N-1$ zero-sequence values between each sequence value of $x[n]$, so that

$$x_p[n] = \begin{cases} x_d\left[\frac{n}{N}\right], & n = 0, \pm N, \pm 2N, \ldots \\ 0, & \text{otherwise} \end{cases}.$$

For exact band-limited interpolation, $H(e^{j\omega})$ is an ideal lowpass filter.

(a) Determine whether or not system A is linear.

(b) Determine whether or not system A is time invariant.

(c) For $X_d(e^{j\omega})$ as sketched in Figure P7.49 and with $N = 3$, sketch $X_p(e^{j\omega})$.

(d) For $N = 3$, $X_d(e^{j\omega})$ as in Figure P7.49, and $H(e^{j\omega})$ appropriately chosen for exact band-limited interpolation, sketch $X(e^{j\omega})$.

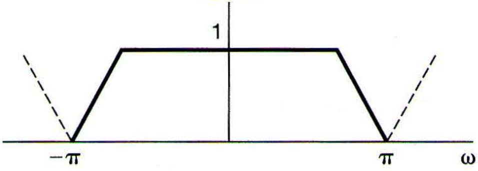


Figure P7.49

7.50. In this problem, we consider the discrete-time counterparts of the zero-order hold and first-order hold, which were discussed for continuous time in Sections 7.1.2 and 7.2.

Let $x[n]$ be a sequence to which discrete-time sampling, as illustrated in Figure 7.31, has been applied. Suppose the conditions of the discrete-time sampling theorem are satisfied; that is, $\omega_s > 2\omega_M$, where ω_s is the sampling frequency and $X(e^{j\omega}) = 0$, $\omega_M < |\omega| \leq \pi$. The original signal $x[n]$ is then exactly recoverable from $x_p[n]$ by ideal lowpass filtering, which, as discussed in Section 7.5, corresponds to band-limited interpolation.

The zero-order hold represents an approximate interpolation whereby every sample value is repeated (or held) $N - 1$ successive times, as illustrated in Figure P7.50(a) for the case of $N = 3$. The first-order hold represents a linear interpolation between samples, as illustrated in the same figure.

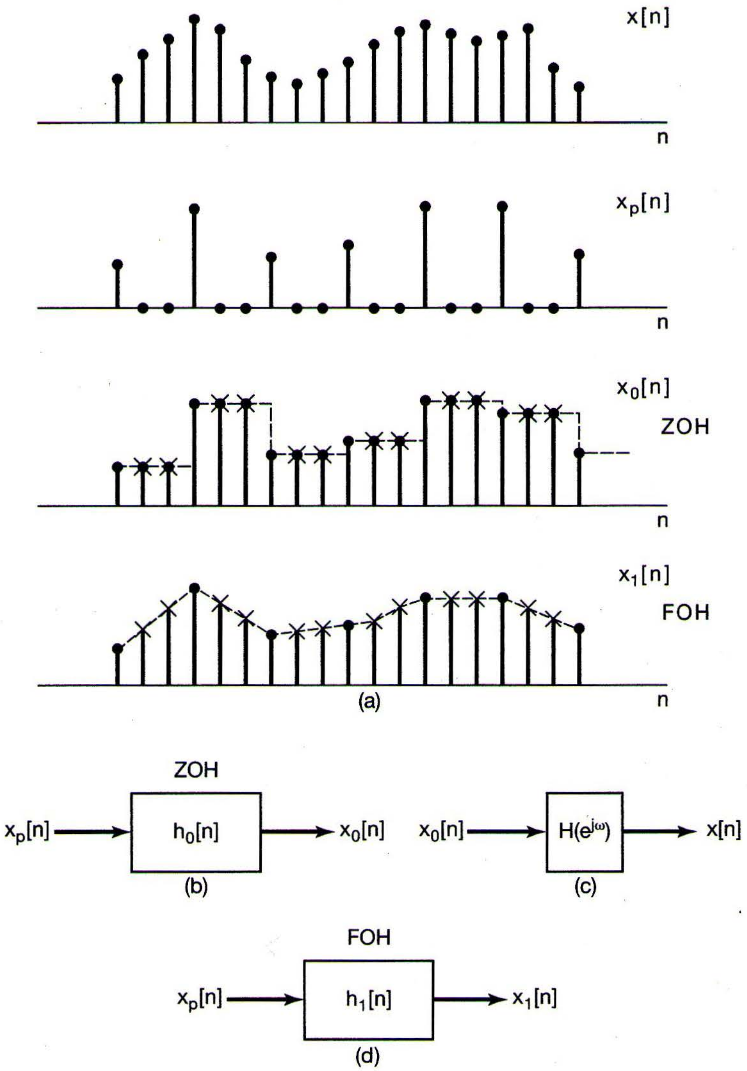


Figure P7.50

(a) The zero-order hold can be represented as an interpolation in the form of eq. (7.47) or, equivalently, the system in Figure P7.50(b). Determine and sketch $h_0[n]$ for the general case of a sampling period N.

(b) $x[n]$ can be exactly recovered from the zero-order-hold sequence $x_0[n]$ using an appropriate LTI filter $H(e^{j\omega})$, as indicated in Figure P7.50(c). Determine and sketch $H(e^{j\omega})$.

(c) The first-order-hold (linear interpolation) can be represented as an interpolation in the form of eq. (7.47) or, equivalently, the system in Figure P7.50(d). Determine and sketch $h_1[n]$ for the general case of a sampling period N.

(d) $x[n]$ can be exactly recovered from the first-order-hold sequence $x_1[n]$ using an appropriate LTI filter with frequency response $H(e^{j\omega})$. Determine and sketch $H(e^{j\omega})$.

7.51. As shown in Figure 7.37 and discussed in Section 7.5.2, the procedure for interpolation or upsampling by an integer factor N can be thought of as a cascade of two operations. For exact band-limited interpolation, the filter $H(e^{j\omega})$ in Figure 7.37 is an ideal lowpass filter. In any specific application, it would be necessary to implement an approximate lowpass filter. In this problem, we explore some useful constraints that are often imposed on the design of these approximate lowpass filters.

(a) Suppose that $H(e^{j\omega})$ is approximated by a zero-phase FIR filter. The filter is to be designed with the constraint that the original sequence values $x_d[n]$ get reproduced *exactly;* that is,

$$x[n] = x_d\left[\frac{n}{L}\right], n = 0, \pm L, \pm 2L, \ldots. \qquad \text{(P7.51–1)}$$

This guarantees that, even though the interpolation between the original sequence values may not be exact, the original values are reproduced exactly in the interpolation. Determine the constraint on the impulse response $h[n]$ of the lowpass filter which guarantees that eq. (P7.51–1) will hold exactly for any sequence $x_d[n]$.

(b) Now suppose that the interpolation is to be carried out with a *linear-phase, causal, symmetric* FIR filter of length N; that is

$$h[n] = 0, n < 0, n > N - 1, \qquad \text{(P7.51–2)}$$

$$H(e^{j\omega}) = H_R(e^{j\omega})e^{-j\alpha\omega}, \qquad \text{(P7.51–3)}$$

where $H_R(e^{j\omega})$ is real. The filter is to be designed with the constraint that the original sequence values $x_d[n]$ get reproduced *exactly,* but with an integer delay α, where α is the negative of the slope of the phase of $H(e^{j\omega})$; that is,

$$x[n] = x_d\left[\frac{n-\alpha}{L}\right], n - \alpha = 0, \pm L, \pm 2L, \ldots \qquad \text{(P7.51–4)}$$

Determine whether this imposes any constraint on whether the filter length N is odd or even.

(c) Again, suppose that the interpolation is to be carried out with a linear-phase, causal, symmetric FIR filter, so that

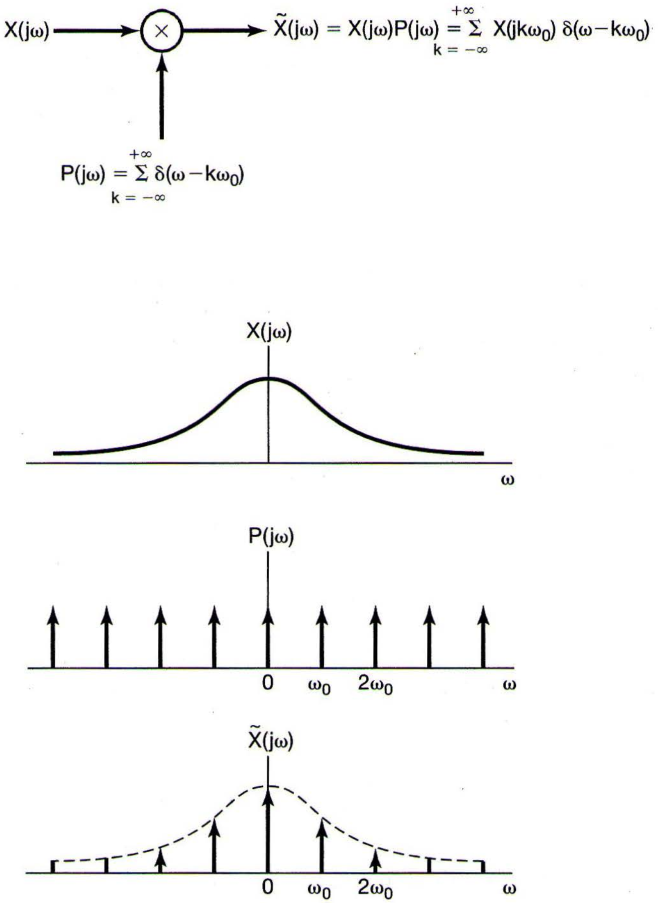
$$H(e^{j\omega}) = H_R(e^{j\omega})e^{-j\beta\omega},$$

where $H_R(e^{j\omega})$ is real. The filter is to be designed with the constraint that the original sequence values $x_d[n]$ get reproduced exactly, but with a delay M that is not necessarily equal to the slope of the phase; that is,

$$x[n] = x_d\left[\frac{n-\alpha}{L}\right], n - M = 0, \pm L, \pm 2L, \ldots.$$

Determine whether this imposes any constraint on whether the filter length N is odd or even.

7.52 In this problem we develop the dual to the time-domain sampling theorem, whereby a time-limited signal can be reconstructed from *frequency-domain* samples. To develop this result, consider the frequency-domain sampling operation in Figure P7.52.



X(jω)
×
$\tilde{X}(j\omega) = X(j\omega)P(j\omega) = \sum_{k=-\infty}^{+\infty} X(jk\omega_0)\,\delta(\omega - k\omega_0)$
$P(j\omega) = \sum_{k=-\infty}^{+\infty} \delta(\omega - k\omega_0)$
X(jω)
ω
P(jω)
0
ω0
2ω0
ω
$\tilde{X}(j\omega)$
0
ω0
2ω0
ω

Figure P7.52

(a) Show that

$$\tilde{x}(t) = x(t) * p(t)$$

where $\tilde{x}(t)$, $x(t)$, and $p(t)$ are the inverse Fourier transforms of $\tilde{X}(j\omega)$, $X(j\omega)$, and $P(j\omega)$, respectively.

(b) Assuming that $x(t)$ is time-limited so that $x(t) = 0$ for $|t| \geq \frac{\pi}{\omega_o}$, show that $x(t)$ can be obtained from $\tilde{x}(t)$ through a "low-time windowing" operation. That is,

$$x(t) = \tilde{x}(t)w(t)$$

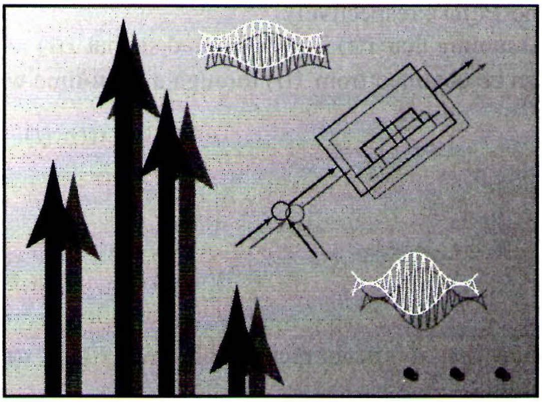
where

$$w(t) = \begin{cases} \omega_0, & |t| \leq \frac{\pi}{\omega_0} \\ 0, & |t| > -\frac{\pi}{\omega_0} \end{cases}$$

(c) Show that $x(t)$ is not recoverable from $\tilde{x}(t)$ if $x(t)$ is not constrained to be zero for $|t| \geq \frac{\pi}{\omega_0}$.

8

COMMUNICATION SYSTEMS



8.0 INTRODUCTION

Communication systems play a key role in our modern world in transmitting information between people, systems, and computers. In general terms, in all communication systems the information at the source is first processed by a transmitter or modulator to change it into a form suitable for transmission over the communication channel. At the receiver, the signal is then recovered through appropriate processing. This processing is required for a variety of reasons. In particular, quite typically, any specific communication channel has associated with it a frequency range over which it is best suited for transmitting a signal and outside of which communication is severely degraded or impossible. For example, the atmosphere will rapidly attenuate signals in the audible frequency range (10 Hz to 20 kHz), whereas it will propagate signals at a higher frequency range over longer distances. Thus, in transmitting audio signals such as speech or music over a communication channel that relies on propagation through the atmosphere, the transmitter first embeds the signal through an appropriate process into another, higher frequency signal.

Many of the concepts and techniques we have developed in the earlier chapters of this text play a central role in the analysis and design of communication systems. As with any concept that is closely tied to a wide variety of important applications, there are a large number of detailed issues to be considered, and, as indicated in the bibliography, there are many excellent texts on the subject. While a full and detailed analysis of communication systems is well beyond the scope of our discussions here, with the background of the previous chapters we are now in a position to introduce some of the basic principles and issues encountered in the design and analysis of these systems.

The general process of embedding an information-bearing signal into a second signal is typically referred to as *modulation*. Extracting the information-bearing signal is known as *demodulation.* As we will see, modulation techniques not only allow us to embed information into signals that can be transmitted effectively, but also make possible the simultaneous transmission of more than one signal with overlapping spectra over the same channel, through a concept referred to as *multiplexing.*

There are a wide variety of modulation methods used in practice, and in this chapter we examine several of the most important of these. One large class of modulation methods relies on the concept of *amplitude modulation* or AM in which the signal we wish to transmit is used to modulate the amplitude of another signal. A very common form of amplitude modulation is *sinusoidal amplitude modulation,* which we explore in some detail in Sections 8.1–8.4 together with the related concepts of frequency-division multiplexing. Another important class of AM systems involves the modulation of the amplitude of a pulsed signal, and in Sections 8.5 and 8.6 we examine this form of modulation as well as the concept of time-division multiplexing. In Section 8.7 we then examine a different form of modulation, namely *sinusoidal frequency modulation* in which the information-bearing signal is used to vary the frequency of a sinusoidal signal.

All of the discussion up through Section 8.7 focuses attention on continuous-time signals, since most transmission media, such as the atmosphere, are best thought of as continuous-time phenomena. Nevertheless, not only is it possible to develop analogous techniques for discrete-time signals, but it is of considerable practical importance to consider modulation concepts involving such signals, and in Section 8.8 we examine some of the basic ideas behind the communication of discrete-time signals.

8.1 COMPLEX EXPONENTIAL AND SINUSOIDAL AMPLITUDE MODULATION

Many communication systems rely on the concept of sinusoidal amplitude modulation, in which a complex exponential or sinusoidal signal $c(t)$ has its amplitude multiplied (modulated) by the information-bearing signal $x(t)$. The signal $x(t)$ is typically referred to as the *modulating signal* and the signal $c(t)$ as the *carrier signal.* The modulated signal $y(t)$ is then the product of these two signals:

$$y(t) = x(t)c(t)$$

As we discussed in Section 8.0, an important objective in modulation is to produce a signal whose frequency range is suitable for transmission over the communication channel to be used. In telephone transmission systems, for example, long-distance transmission is often accomplished over microwave or satellite links. The individual voice signals are in the frequency range 200 Hz to 4 kHz, whereas a microwave link requires signals in the range 300 megahertz (MHz) to 300 gigahertz (GHz), and communication satellite links operate in the frequency range from a few hundred MHz to over 40 GHz. Thus, for transmission over these channels, the information in a voice signal must be shifted into these higher ranges of frequency. As we will see in this section, sinusoidal amplitude modulation achieves such a shift in frequency in a very simple manner.

8.1.1 Amplitude Modulation with a Complex Exponential Carrier

There are two common forms of sinusoidal amplitude modulation, one in which the carrier signal is a complex exponential of the form

$$c(t) = e^{j(\omega_c t+\theta_c)} \tag{8.1}$$

and the second in which the carrier signal is sinusoidal and of the form

$$c(t) = \cos(\omega_c t + \theta_c). \tag{8.2}$$

In both cases, the frequency ω_c is referred to as the *carrier frequency.* Let us consider first the case of a complex exponential carrier, and for convenience, let us choose $\theta_c = 0$, so that the modulated signal is

$$y(t) = x(t)e^{j\omega_c t}. \tag{8.3}$$

From the multiplication property (Section 4.5), and with $X(j\omega)$, $Y(j\omega)$, and $C(j\omega)$ denoting the Fourier transforms of $x(t)$, $y(t)$, and $c(t)$, respectively,

$$Y(j\omega) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} X(j\theta)C(j(\omega - \theta))d\theta. \tag{8.4}$$

For $c(t)$ a complex exponential as given in eq. (8.1),

$$C(j\omega) = 2\pi\delta(\omega - \omega_c), \tag{8.5}$$

and hence,

$$Y(j\omega) = X(j\omega - j\omega_c). \tag{8.6}$$

Thus, the spectrum of the modulated output $y(t)$ is simply that of the input, shifted in frequency by an amount equal to the carrier frequency ω_c. For example, with $X(j\omega)$ band limited with highest frequency ω_M (and bandwidth $2\omega_M$), as depicted in Figure 8.1(a), the output spectrum $Y(j\omega)$ is that shown in Figure 8.1(c).

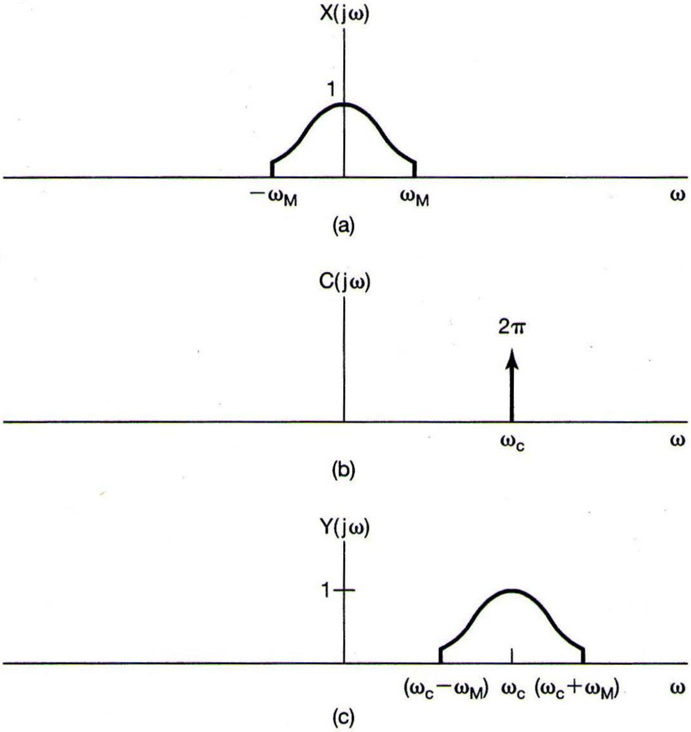


Figure 8.1 Effect in the frequency domain of amplitude modulation with a complex exponential carrier: (a) spectrum of modulating signal $x(t)$; (b) spectrum of carrier $c(t) = e^{j\omega_c t}$; (c) spectrum of amplitude-modulated signal $y(t) = x(t)e^{j\omega_c t}$.

From eq. (8.3), it is clear that $x(t)$ can be recovered from the modulated signal $y(t)$ by multiplying by the complex exponential $e^{-j\omega_c t}$; that is,

$$x(t) = y(t)e^{-j\omega_c t}. \tag{8.7}$$

In the frequency domain, this has the effect of shifting the spectrum of the modulated signal back to its original position on the frequency axis. The process of recovering the original signal from the modulated signal is referred to as *demodulation,* a topic we discuss at more length in Section 8.2.

Since $e^{j\omega_c t}$ is a complex signal, eq. (8.3) can be rewritten as

$$y(t) = x(t)\cos\omega_c t + jx(t)\sin\omega_c t. \tag{8.8}$$

Implementation of eq. (8.7) or (8.8) with $x(t)$ real utilizes two separate multipliers and two sinusoidal carrier signals that have a phase difference of $\pi/2$, as depicted in Figure 8.2 for $c(t)$ given by eq. (8.1). In Section 8.4 we give an example of one of the applications in which there are particular advantages to using a system, such as in Figure 8.2, employing two sinusoidal carriers with a phase difference of $\pi/2$.

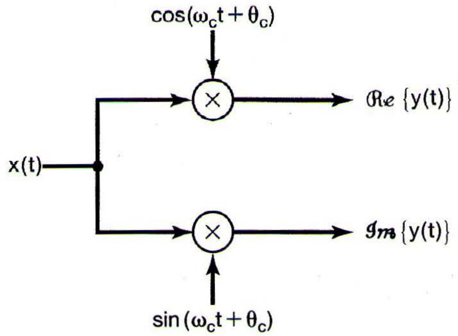


Figure 8.2 Implementation of amplitude modulation with a complex exponential carrier $c(t) = e^{j(\omega_c t + \theta_c)}$.

8.1.2 Amplitude Modulation with a Sinusoidal Carrier

In many situations, using a sinusoidal carrier of the form of eq. (8.2) is often simpler than and equally as effective as using a complex exponential carrier. In effect, using a sinusoidal carrier corresponds to retaining only the real or imaginary part of the output of Figure 8.2. A system that uses a sinusoidal carrier is depicted in Figure 8.3.

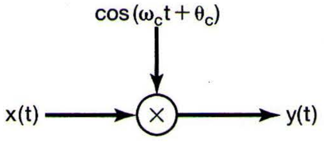


Figure 8.3 Amplitude modulation with a sinusoidal carrier.

The effect of amplitude modulation with a sinusoidal carrier in the form of eq. (8.2) can be analyzed in a manner identical to that in the preceding subsection. Again, for convenience we choose $\theta_c = 0$. In this case, the spectrum of the carrier signal is

$$C(j\omega) = \pi[\delta(\omega - \omega_c) + \delta(\omega + \omega_c)], \tag{8.9}$$

and thus, from eq. (8.4),

$$Y(j\omega) = \frac{1}{2}[X(j\omega - j\omega_c) + X(j\omega + j\omega_c)]. \tag{8.10}$$

With $X(j\omega)$ as depicted in Figure 8.4(a), the spectrum of $y(t)$ is that shown in Figure 8.4(c). Note that there is now a replication of the spectrum of the original signal, centered around both $+\omega_c$ and $-\omega_c$. As a consequence, $x(t)$ is recoverable from $y(t)$ only if $\omega_c > \omega_M$, since otherwise the two replications will overlap in frequency. This is in contrast to the case of a complex exponential carrier, for which a replication of the spectrum of the original signal is centered only around ω_c. Specifically, as we saw in Section 8.1.1, in the case of amplitude modulation with a complex exponential carrier, $x(t)$ can always be recovered from $y(t)$ for any choice of ω_c by shifting the spectrum back to its original location by multiplying by $e^{-j\omega_c t}$, as in eq. (8.7). With a sinusoidal carrier, on the other hand, as we see from Figure 8.4, if $\omega_c < \omega_M$, then there will be an overlap between the two replications of $X(j\omega)$. For example, Figure 8.5 depicts $Y(j\omega)$ for $\omega_c = \omega_M/2$. Clearly, the spectrum of $x(t)$ is no longer replicated in $Y(j\omega)$, and thus, it may no longer be possible to recover $x(t)$ from $y(t)$.

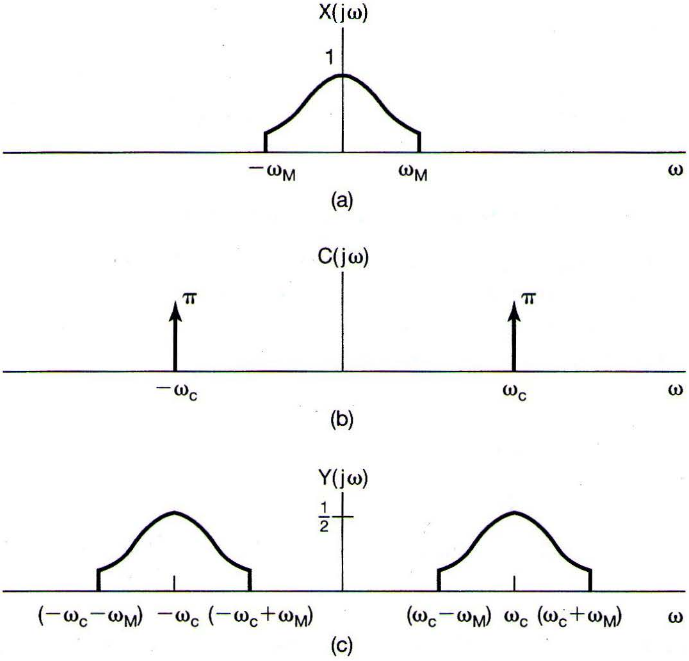


Figure 8.4 Effect in the frequency domain of amplitude modulation with a sinusoidal carrier: (a) spectrum of modulating signal $x(t)$; (b) spectrum of carrier $c(t) = \cos \omega_c t$; (c) spectrum of amplitude-modulated signal.

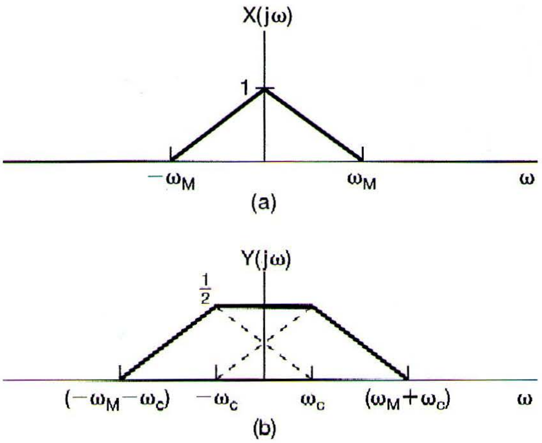


Figure 8.5 Sinusoidal amplitude modulation with carrier $\cos \omega_c t$ for which $\omega_c = \omega_M/2$: (a) spectrum of modulating signal; (b) spectrum of modulated signal.

8.2 DEMODULATION FOR SINUSOIDAL AM

At the receiver in a communication system, the information-bearing signal $x(t)$ is recovered through demodulation. In this section, we examine the process of demodulation for sinusoidal amplitude modulation, as introduced in the previous section. There are two commonly used methods for demodulation, each with its own advantages and disadvantages. In Section 8.2.1 we discuss the first of these, a process referred to as *synchronous demodulation,* in which the transmitter and receiver are synchronized in phase. In Section 8.2.2, we describe an alternative method referred to as *asynchronous demodulation.*

8.2.1 Synchronous Demodulation

Assuming that $\omega_c > \omega_M$, demodulation of a signal that was modulated with a sinusoidal carrier is relatively straightforward. Specifically, consider the signal

$$y(t) = x(t) \cos \omega_c t. \tag{8.11}$$

As was suggested in Example 4.21, the original signal can be recovered by modulating $y(t)$ with the same sinusoidal carrier and applying a lowpass filter to the result. To see this, consider

$$w(t) = y(t) \cos \omega_c t. \tag{8.12}$$

Figure 8.6 shows the spectra of $y(t)$ and $w(t)$, and we observe that $x(t)$ can be recovered from $w(t)$ by applying an ideal lowpass filter with a gain of 2 and a cutoff frequency that is greater than ω_M and less than $2\omega_c - \omega_M$. The frequency response of the lowpass filter is indicated by the dashed line in Figure 8.6(c).

The basis for using eq. (8.12) and a lowpass filter to demodulate $y(t)$ can also be seen algebraically. From eqs. (8.11) and (8.12), it follows that

$$w(t) = x(t) \cos^2 \omega_c t,$$

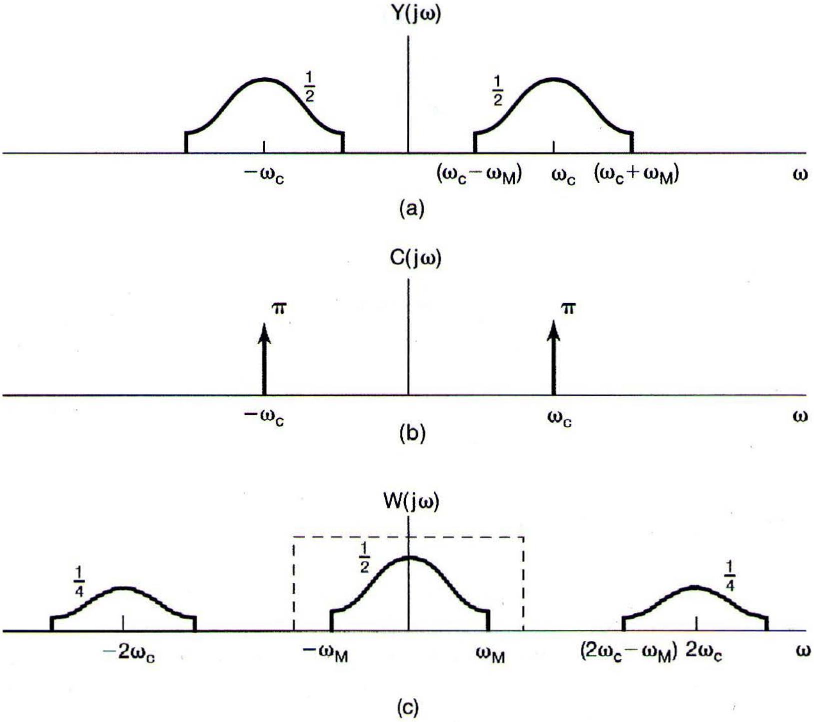


Figure 8.6 Demodulation of an amplitude-modulated signal with a sinusoidal carrier: (a) spectrum of modulated signal; (b) spectrum of carrier signal; (c) spectrum of modulated signal multiplied by the carrier. The dashed line indicates the frequency response of a lowpass filter used to extract the demodulated signal.

or, using the trigonometric identity

$$\cos^2 \omega_c t = \frac{1}{2} + \frac{1}{2}\cos 2\omega_c t,$$

we can rewrite $w(t)$ as

$$w(t) = \frac{1}{2}x(t) + \frac{1}{2}x(t)\cos 2\omega_c t. \tag{8.13}$$

Thus, $w(t)$ consists of the sum of two terms, namely one-half the original signal and one-half the original signal modulated with a sinusoidal carrier at twice the original carrier frequency ω_c. Both of these terms are apparent in the spectrum shown in Figure 8.6(c). Applying the lowpass filter to $w(t)$ corresponds to retaining the first term on the right-hand side of eq. (8.13) and eliminating the second term.

The overall system for amplitude modulation and demodulation using a complex exponential carrier is depicted in Figure 8.7, and the overall system for modulation and demodulation using a sinusoidal carrier is depicted in Figure 8.8. In these figures, we have indicated the more general case in which, for both the complex exponential and the sinusoidal carrier, a carrier phase θ_c is included. The modification of the preceding analysis so as to include θ_c is straightforward and is considered in Problem 8.21.

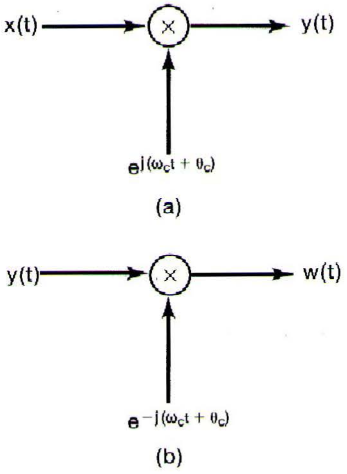


Figure 8.7 System for amplitude modulation and demodulation using a complex exponential carrier: (a) modulation; (b) demodulation.

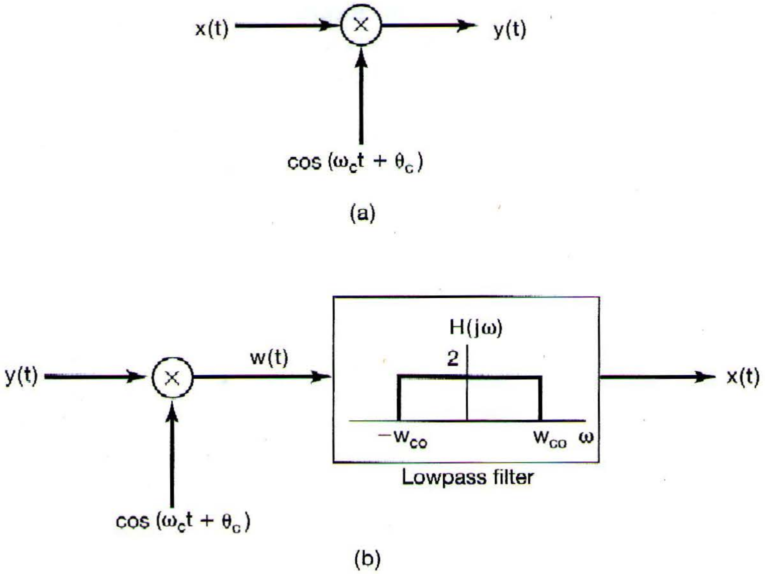


Figure 8.8 Amplitude modulation and demodulation with a sinusoidal carrier: (a) modulation system; (b) demodulation system. The lowpass filter cutoff frequency W_{co} is greater than ω_M and less than $2\omega_c - \omega_M$.

In the systems of Figures 8.7 and 8.8, the demodulating signal is assumed to be synchronized in phase with the modulating signal, and consequently the process is referred to as *synchronous demodulation*. Suppose, however, that the modulator and demodulator are not synchronized in phase. For the case of the complex exponential carrier, with θ_c denoting the phase of the modulating carrier and ϕ_c the phase of the demodulating carrier,

$$y(t) = e^{j(\omega_c t+\theta_c)}x(t), \tag{8.14}$$

$$w(t) = e^{-j(\omega_c t+\phi_c)}y(t), \tag{8.15}$$

and consequently,

$$w(t) = e^{j(\theta_c-\phi_c)}x(t). \tag{8.16}$$

Thus, if $\theta_c \neq \phi_c$, $w(t)$ will have a complex amplitude factor. For the particular case in which $x(t)$ is positive, $x(t) = |w(t)|$, and thus $x(t)$ can be recovered by taking the magnitude of the demodulated signal.

For the sinusoidal carrier, again let θ_c and ϕ_c denote the phases of the modulating and demodulating carriers, respectively, as indicated in Figure 8.9. The input to the lowpass filter is now

$$w(t) = x(t)\cos(\omega_c t + \theta_c)\cos(\omega_c t + \phi_c), \tag{8.17}$$

or, using the trigonometric identity

$$\cos(\omega_c t + \theta_c)\cos(\omega_c t + \phi_c) = \frac{1}{2}\cos(\theta_c - \phi_c) + \frac{1}{2}\cos(2\omega_c t + \theta_c + \phi_c), \tag{8.18}$$

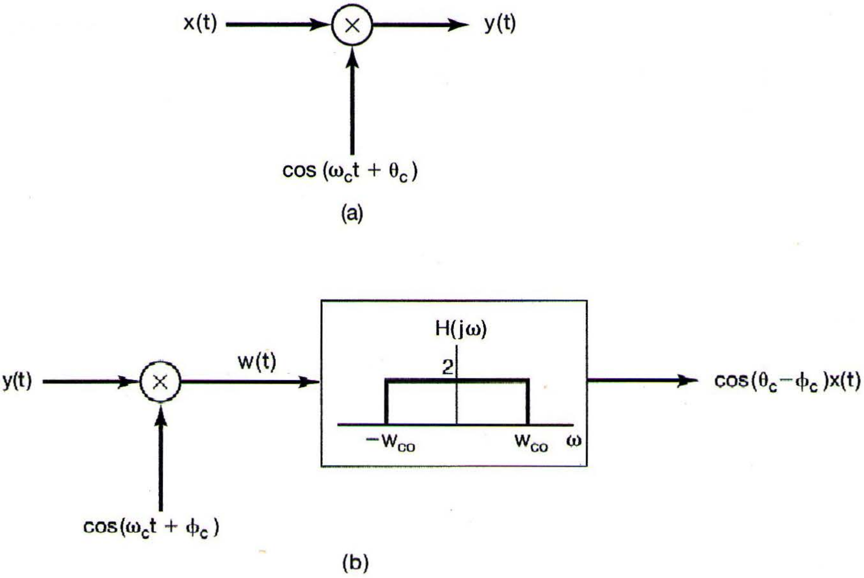


Figure 8.9 Sinusoidal amplitude modulation and demodulation system for which the carrier signals and the modulator and demodulator are not synchronized: (a) modulator; (b) demodulator.

we have

$$w(t) = \frac{1}{2}\cos(\theta_c - \phi_c)x(t) + \frac{1}{2}x(t)\cos(2\omega_c t + \theta_c + \phi_c), \tag{8.19}$$

and the output of the lowpass filter is then $x(t)$ multiplied by the amplitude factor $\cos(\theta_c - \phi_c)$. If the oscillators in the modulator and demodulator are in phase, $\theta_c = \phi_c$, and the output of the lowpass filter is $x(t)$. On the other hand, if these oscillators have a phase difference of $\pi/2$, the output will be zero. In general, for a maximum output signal, the oscillators should be in phase. Of even more importance, the phase relation between the two oscillators must be maintained over time, so that the amplitude factor $\cos(\theta_c - \phi_c)$ does not vary. This requires careful synchronization between the modulator and the demodulator, which is often difficult, particularly when they are geographically separated, as is typical in a communication system. The corresponding effects of, and the need for, synchronization not only between the phase of the modulator and demodulator, but between the frequencies of the carrier signals used in both, are explored in detail in Problem 8.23.

8.2.2 Asynchronous Demodulation

In many systems that employ sinusoidal amplitude modulation, an alternative demodulation procedure referred to as *asynchronous demodulation* is commonly used. Asynchronous demodulation avoids the need for synchronization between the modulator and demodulator. In particular, suppose that $x(t)$ is always positive and that the carrier frequency ω_c is much higher than ω_M, the highest frequency in the modulating signal. The modulated signal $y(t)$ will then have the general form illustrated in Figure 8.10.

In particular, the *envelope* of $y(t)$—that is, a smooth curve connecting the peaks in $y(t)$—would appear to be a reasonable approximation to $x(t)$. Thus, $x(t)$ could be approximately recovered through the use of a system that tracks these peaks to extract the envelope. Such a system is referred to as an *envelope detector.* One example of a simple circuit that acts as an envelope detector is shown in Figure 8.11(a). This circuit is generally followed by a lowpass filter to reduce the variations at the carrier frequency, which are evident in Figure 8.11(b) and which will generally be present in the output of an envelope detector of the type indicated in Figure 8.11(a).

The two basic assumptions required for asynchronous demodulation are that $x(t)$ be positive and that $x(t)$ vary slowly compared to ω_c, so that the envelope is easily tracked. The second condition is satisfied, for example, in audio transmission over a radio-frequency (RF) channel, where the highest frequency present in $x(t)$ is typically 15 to 20 kHz and $\omega_c/2\pi$ is in the range 500 kHz to 2 MHz. The first condition, that $x(t)$ be positive, can be satisfied by simply adding an appropriate constant value to $x(t)$ or, equivalently, by a simple change in the modulator, as shown in Figure 8.12. The output of the envelope detector then approximates $x(t) + A$, from which $x(t)$ is easily obtained.

To use the envelope detector for demodulation, we require that A be sufficiently large so that $x(t) + A$ is positive. Let K denote the maximum amplitude of $x(t)$; that is, $|x(t)| \leq K$. For $x(t) + A$ to be positive, we require that $A > K$. The ratio K/A is commonly referred to as the *modulation index m.* Expressed in percent, it is referred to as the *percent modulation.* An illustration of the output of the modulator of Figure 8.12 for $x(t)$ sinusoidal and for $m = 0.5$ (50% modulation) and $m = 1.0$ (100% modulation), is shown in Figure 8.13.

In Figure 8.14, we show a comparison of the spectra associated with the modulated signal when synchronous demodulation and when asynchronous demodulation are used. We note in particular that the output of the modulator for the asynchronous system in Figure 8.12 has an additional component $A \cos \omega_c t$ that is neither present nor necessary in the synchronous system. This is represented in the spectrum of Figure 8.14(c) by the presence of impulses at $+\omega_c$ and $-\omega_c$. For a fixed maximum amplitude K of the modulating signal, as A is decreased the relative amount of carrier present in the modulated output decreases. Since the carrier component in the output contains no information, its presence

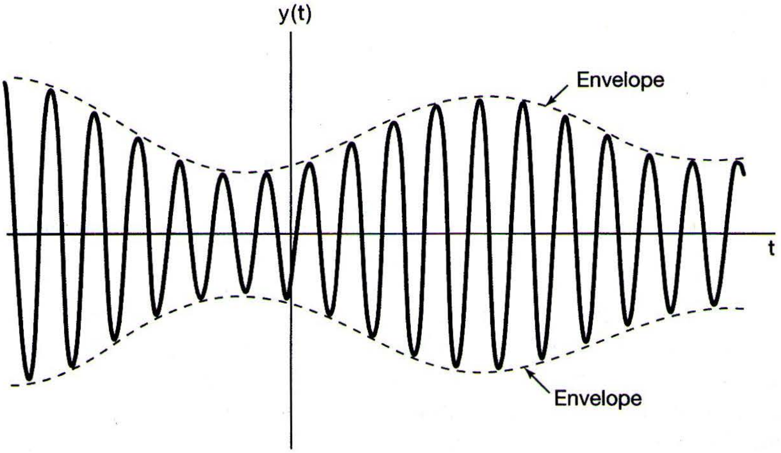
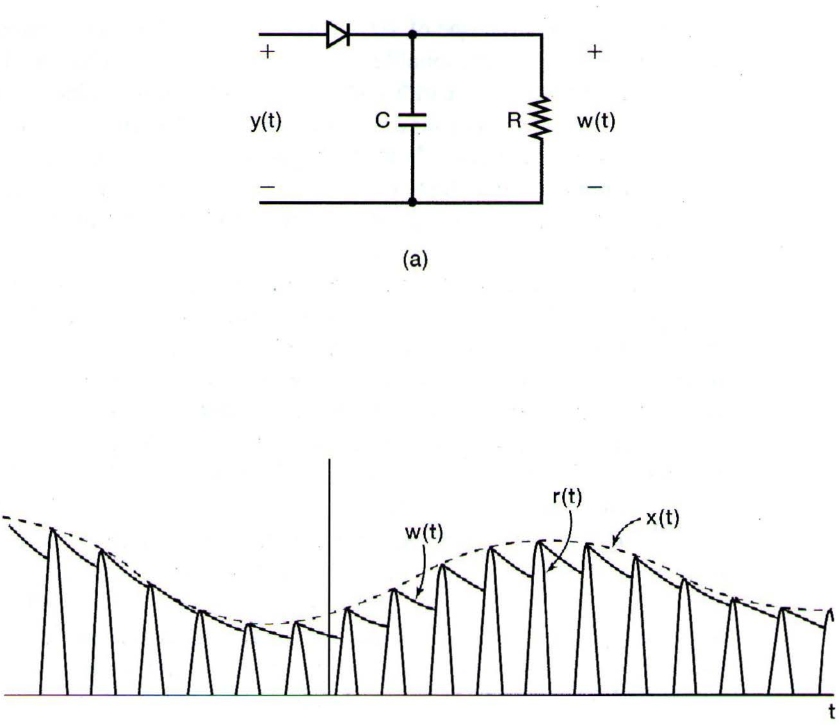


Figure 8.10 Amplitude-modulated signal for which the modulating signal is positive. The dashed curve represents the envelope of the modulated signal.



(a)

(b)

Figure 8.11 Demodulation by envelope detection: (a) circuit for envelope detection using half-wave rectification; (b) waveforms associated with the envelope detector in (a): $r(t)$ is the half-wave rectified signal, $x(t)$ is the true envelope, and $w(t)$ is the envelope obtained from the circuit in (a). The relationship between $x(t)$ and $w(t)$ has been exaggerated in (b) for purposes of illustration. In a practical asynchronous demodulation system, $w(t)$ would typically be a much closer approximation to $x(t)$ than depicted here.

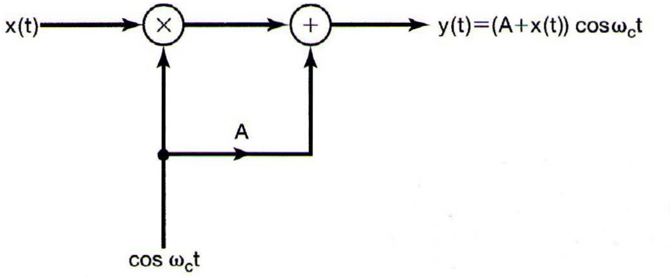


Figure 8.12 Modulator for an asynchronous modulation-demodulation system.

represents an inefficiency—for example, in the amount of power required to transmit the modulated signal—and thus, in one sense it is desirable to make the ratio K/A—i.e., the modulation index m—as large as possible. On the other hand, the ability of a simple envelope detector such as that in Figure 8.11 to follow the envelope and thus extract $x(t)$ improves as the modulation index decreases. Hence, there is a trade-off between the effi-

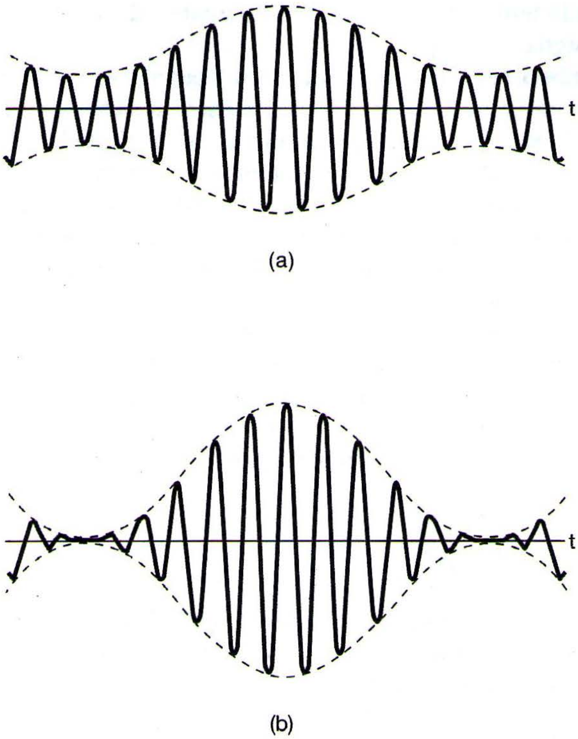


Figure 8.13 Output of the amplitude modulation system of Figure 8.12: (a) modulation index $m = 0.5$; (b) modulation index $m = 1.0$.

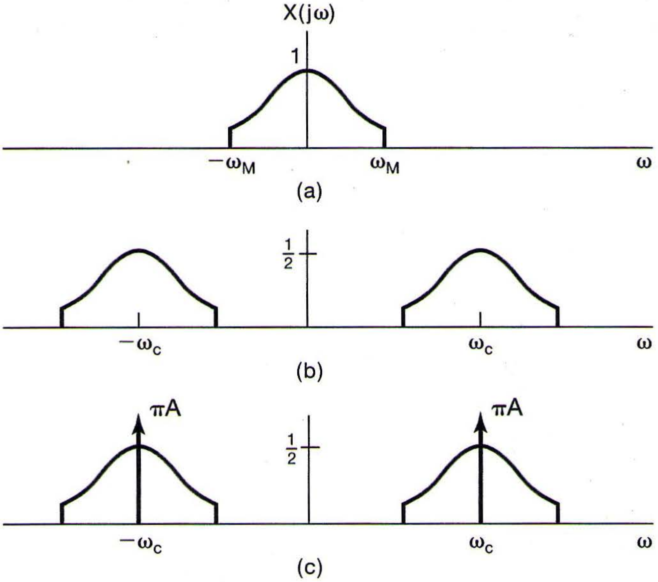


Figure 8.14 Comparison of spectra for synchronous and asynchronous sinusoidal amplitude modulation systems: (a) spectrum of modulating signal; (b) spectrum of $x(t) \cos \omega_c t$ representing modulated signal in a synchronous system; (c) spectrum of $[x(t) + A] \cos \omega_c t$ representing modulated signal in an asynchronous system.

ciency of the system in terms of the power in the output of the modulator and the quality of the demodulated signal.

There are a number of advantages and disadvantages to the asynchronous modulation-demodulation system of Figures 8.11 and 8.12, compared with the synchronous system of Figure 8.8. The synchronous system requires a more sophisticated demodulator because the oscillator in the demodulator must be synchronized with the oscillator in the modulator, both in phase and in frequency. On the other hand, the asynchronous modulator in general requires transmitting more power than the synchronous modulator, since, for the envelope detector to operate properly, the envelope must be positive, or equivalently, there must be a carrier component present in the transmitted signal. This is often preferable in cases such as that associated with public radio broadcasting, in which it is desirable to mass-produce large numbers of receivers (demodulators) at moderate cost. The additional cost in transmitted power is then offset by the savings in cost for the receiver. On the other hand, in situations in which transmitter power requirements are at a premium, as in satellite communication, the cost of implementing a more sophisticated synchronous receiver is warranted.

8.3 FREQUENCY-DIVISION MULTIPLEXING

Many systems used for transmitting signals provide more bandwidth than is required for any one signal. For example, a typical microwave link has a total bandwidth of several gigahertz, which is considerably greater than the bandwidth required for one voice channel. If the individual voice signals, which are overlapping in frequency, have their frequency content shifted by means of sinusoidal amplitude modulation so that the spectra of the modulated signals no longer overlap, they can be transmitted *simultaneously* over a single wideband channel. The resulting concept is referred to as *frequency-division multiplexing* (FDM). Frequency-division multiplexing using a sinusoidal carrier is illustrated in Figure 8.15. The individual signals to be transmitted are assumed to be band limited and

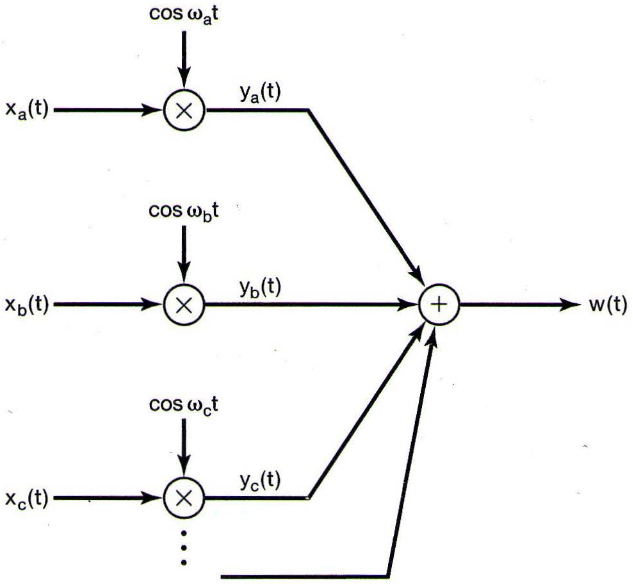


Figure 8.15 Frequency-division multiplexing using sinusoidal amplitude modulation.

are modulated with different carrier frequencies. The modulated signals are then summed and transmitted simultaneously over the same communication channel. The spectra of the individual subchannels and the composite multiplexed signal are illustrated in Figure 8.16. Through this multiplexing process, the individual input signals are allocated distinct segments of the frequency band. To recover the individual channels in the demultiplexing process requires two basic steps: bandpass filtering to extract the modulated signal corresponding to a specific channel, followed by demodulation to recover the original signal. This is illustrated in Figure 8.17 to recover channel *a*, where, for purposes of illustration, synchronous demodulation is assumed.

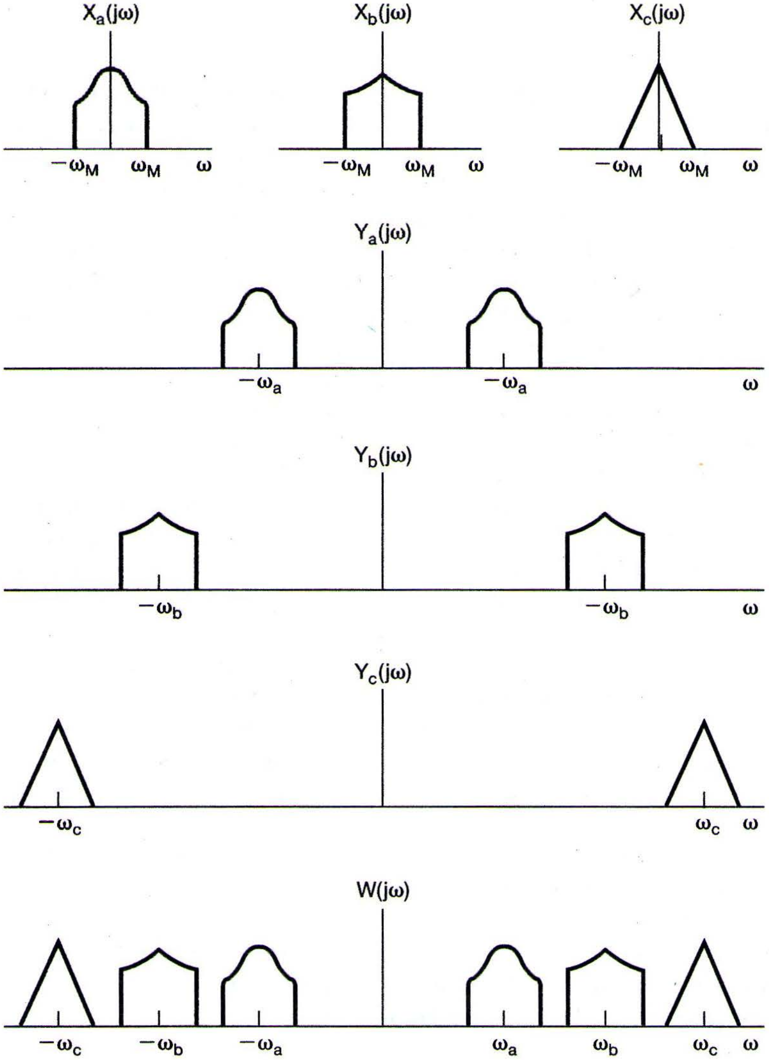

$X_a(j\omega)$
$-\omega_M$
ω_M
ω
$X_b(j\omega)$
$-\omega_M$
ω_M
ω
$X_c(j\omega)$
$-\omega_M$
ω_M
ω
$Y_a(j\omega)$
$-\omega_a$
$-\omega_a$
ω
$Y_b(j\omega)$
$-\omega_b$
$-\omega_b$
ω
$Y_c(j\omega)$
$-\omega_c$
ω_c
ω
$W(j\omega)$
$-\omega_c$
$-\omega_b$
$-\omega_a$
ω_a
ω_b
ω_c
ω

Figure 8.16 Spectra associated with the frequency-division multiplexing system of Figure 8.15.

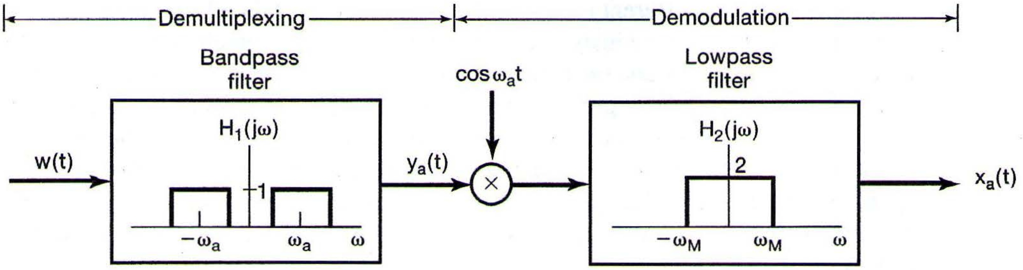


Figure 8.17 Demultiplexing and demodulation for a frequency-division multiplexed signal.

Telephone communication is one important application of frequency-division multiplexing. Another is the transmission of signals through the atmosphere in the RF band. In the United States, the use of radio frequencies for transmitting signals over the range 10 kHz to 275 GHz is controlled by the Federal Communications Commission, and different portions of the range are allocated for different purposes. The current allocation of frequencies is shown in Figure 8.18. As indicated, the frequency range in the neighborhood of 1 MHz is assigned to the AM broadcast band, where *AM* refers specifically to the use of sinusoidal amplitude modulation. Individual AM radio stations are assigned specific frequencies within the AM band, and thus, many stations can broadcast simultaneously through this use of frequency-division multiplexing. In principle, at the receiver, an individual radio station can be selected by demultiplexing and demodulating, as illustrated in Figure 8.17. The tuning dial on the receiver would then control both the center frequency of the bandpass filter and the frequency of the demodulating oscillator. In fact, for public broadcasting, asynchronous modulation and demodulation are used to simplify the receiver and reduce its cost. Furthermore, the demultiplexing in Figure 8.17 requires a sharp cutoff bandpass filter with variable center frequency. Variable frequency-selective filters are difficult to implement, and consequently, a fixed filter is implemented instead, and an intermediate stage of modulation and filtering [referred to in a radio receiver as the *intermediate-frequency* (IF) stage] is used. The use of modulation to slide the spectrum of the signal past a fixed bandpass filter replaces the use of a variable bandpass filter in a manner similar to the procedure discussed in Section 4.5.1. This basic procedure is incorporated into typical home AM radio receivers. Some of the more detailed issues involved are considered in Problem 8.36.

As illustrated in Figure 8.16, in the frequency-division multiplexing system of Figure 8.15 the spectrum of each individual signal is replicated at both positive and negative frequencies, and thus the modulated signal occupies twice the bandwidth of the original. This represents an inefficient use of bandwidth. In the next section we consider an alternative form of sinusoidal amplitude modulation, which leads to more efficient use of bandwidth at the cost of a more complicated modulation system.

Frequency range	Designation	Typical uses	Propagation method	Channel features
30–300 Hz	ELF (extremely low frequency)	Macrowave, submarine communication	Megametric waves	Penetration of conducting earth and seawater
0.3–3 kHz	VF (voice frequency)	Data terminals, telephony	Copper wire	
3–30 kHz	VLF (very low frequency)	Navigation, telephone, telegraph, frequency and timing standards	Surface ducting (ground wave)	Low attenuation, little fading, extremely stable phase and frequency, large antennas
30–300 kHz	LF (low frequency)	Industrial (power line) communication, aeronautical and maritime long-range navigation, radio beacons	Mostly surface ducting	Slight fading, high atmospheric pulse
0.3–3 MHz	MF (medium frequency)	Mobile, AM broadcasting, amateur, public safety	Ducting and ionospheric reflection (sky wave)	Increased fading, but reliable
3–30 MHz	HF (high frequency)	Military communication, aeronautical mobile, international fixed, amateur and citizen's band, industrial	Ionospheric reflecting sky wave, 50–400 km layer altitudes	Intermittent and frequency-selective fading, multipath
30–300 MHz	VHF (very high frequency)	FM and TV broadcast, land transportation (taxis, buses, railroad)	Sky wave (ionospheric and tropospheric scatter)	Fading, scattering, and multipath
0.3–3 GHz	UHF (ultra high frequency)	UHF TV, space telemetry, radar, military	Transhorizon tropospheric scatter and line-of-sight relaying	
3–30 GHz	SHF (super high frequency)	Satellite and space communication, common carrier (CC), microwave	Line-of-sight ionosphere penetration	Ionospheric penetration, extraterrestrial noise, high directly
30–300 GHz	EHF (extremely high frequency)	Experimental, government, radio astronomy	Line of sight	Water vapor and oxygen absorption
10^3–10^7 GHz	Infrared, visible light, ultraviolet	Optical communications	Line of sight	

Figure 8.18 Allocation of frequencies in the RF spectrum.

8.4 SINGLE-SIDEBAND SINUSOIDAL AMPLITUDE MODULATION

For the sinusoidal amplitude modulation systems discussed in Section 8.1, the total bandwidth of the original signal $x(t)$ is $2\omega_M$, including both positive and negative frequencies, where ω_M is the highest frequency present in $x(t)$. With the use of a complex exponential carrier, the spectrum is translated to ω_c, and the total width of the frequency band over which there is energy from the signal is still $2\omega_M$, although the modulated signal is now complex. With a sinusoidal carrier, on the other hand, the spectrum of the signal is shifted to $+\omega_c$ *and* $-\omega_c$, and thus, twice the bandwidth is required. This suggests that there is a basic redundancy in the modulated signal with a sinusoidal carrier. Using a technique referred to as *single-sideband modulation,* we can remove the redundancy.

The spectrum of $x(t)$ is illustrated in Figure 8.19(a), in which we have shaded the positive and negative frequency components differently to distinguish them. The spectrum in Figure 8.19(b) results from modulation with a sinusoidal carrier, where we identify an upper and lower sideband for the portion of the spectrum centered at $+\omega_c$ and that centered at $-\omega_c$. Comparing Figures 8.19(a) and (b), we see that $X(j\omega)$ can be recovered if only the upper sidebands at positive and negative frequencies are retained, or alternatively, if only the lower sidebands at positive and negative frequencies are retained. The resulting spectrum if only the upper sidebands are retained is shown in Figure 8.19(c), and the resulting spectrum if only the lower sidebands are retained is shown in Figure 8.19(d). The conversion of $x(t)$ to the form corresponding to Figure 8.19(c) or (d) is referred to as *single-sideband modulation* (SSB), in contrast to the *double-sideband modulation* (DSB) of Figure 8.19(b), in which both sidebands are retained.

There are several methods by which the single-sideband signal can be obtained. One is to apply a sharp cutoff bandpass or highpass filter to the double-sideband signal of Figure 8.19(b), as illustrated in Figure 8.20, to remove the unwanted sideband. Another is to use a procedure that utilizes phase shifting. Figure 8.21 depicts a system designed

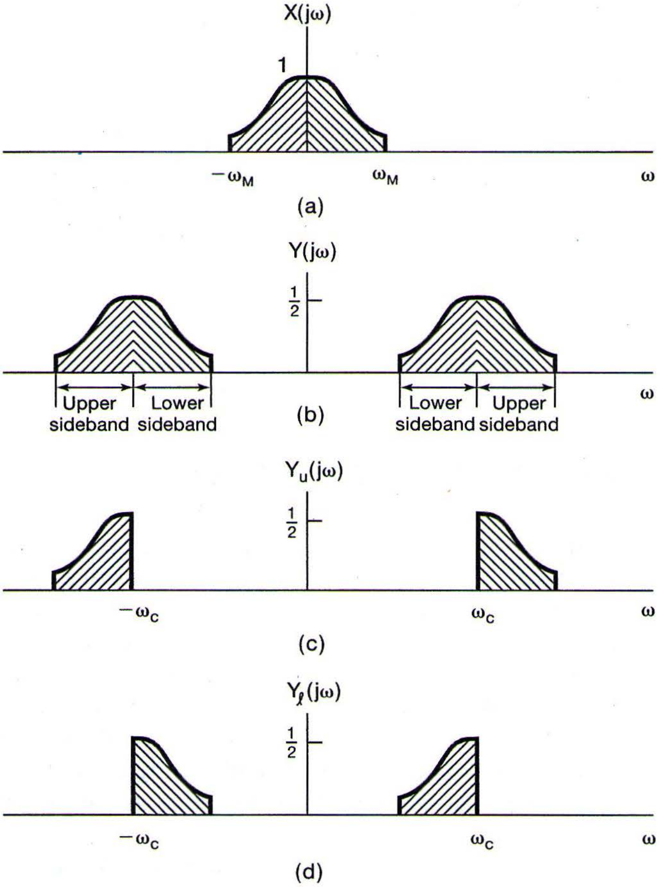


Figure 8.19 Double- and single-sideband modulation: (a) spectrum of modulating signal; (b) spectrum after modulation with a sinusoidal carrier; (c) spectrum with only the upper sidebands; (d) spectrum with only the lower sidebands.

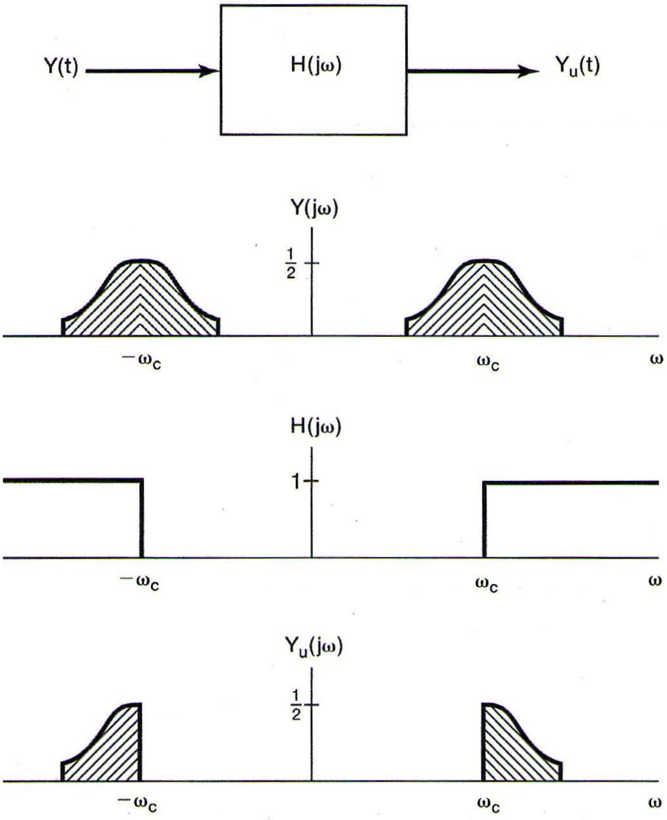


Figure 8.20 System for retaining the upper sidebands using ideal highpass filtering.

to retain the lower sidebands. The system $H(j\omega)$ in the figure is referred to as a "90° phase-shift network," for which the frequency response is of the form

$$H(j\omega) = \begin{cases} -j, & \omega > 0 \\ j, & \omega < 0 \end{cases}. \tag{8.20}$$

The spectra of $x(t)$, $y_1(t) = x(t)\cos\omega_c t$, $y_2(t) = x_p(t)\sin\omega_c t$, and $y(t)$ are illustrated in Figure 8.22. As is examined in Problem 8.28, to retain the upper sidebands instead of the lower sidebands, the phase characteristic of $H(j\omega)$ is reversed so that

$$H(j\omega) = \begin{cases} j, & \omega > 0 \\ -j, & \omega < 0 \end{cases}. \tag{8.21}$$

As is explored in Problem 8.29, synchronous demodulation of single-sideband systems can be accomplished in a manner identical to synchronous demodulation of double-sideband systems. The price paid for the increased efficiency of single-sideband systems is added complexity in the modulator.

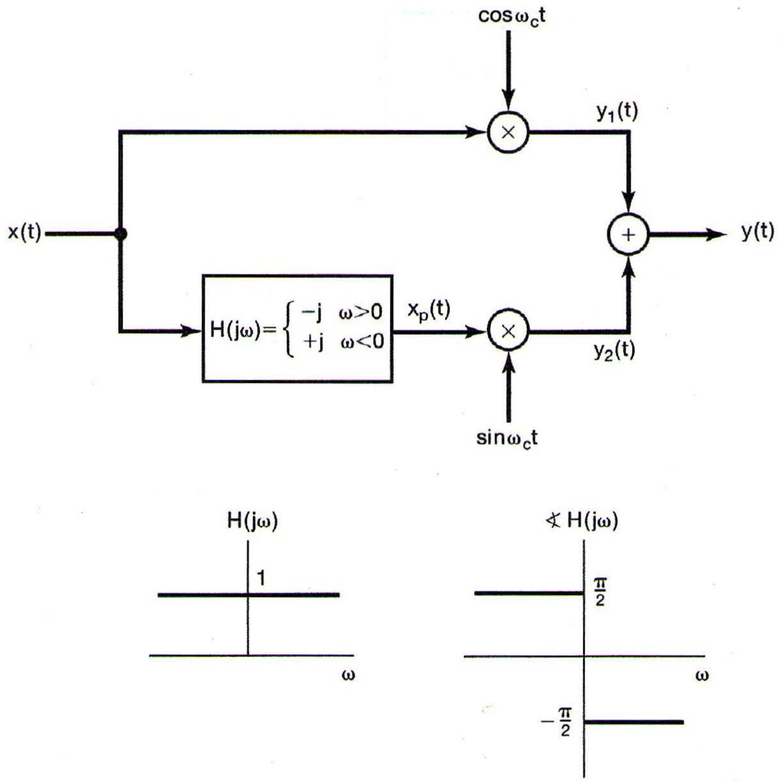


Figure 8.21 System for single-sideband amplitude modulation, using a 90° phase-shift network, in which only the lower sidebands are retained.

In summary, in Sections 8.1 through 8.4 we have seen a number of variations of complex exponential and sinusoidal amplitude modulation. With asynchronous demodulation, discussed in Section 8.2.2, a constant must be added to the modulating signal so that it is positive. This results in the presence of the carrier signal as a component in the modulated output, requiring more power for transmission, but resulting in a simpler demodulator than is required in a synchronous system. Alternatively, only the upper or lower sidebands in the modulated output may be retained, which makes more efficient use of bandwidth and transmitter power, but requires a more sophisticated modulator. Sinusoidal amplitude modulation with both sidebands and the presence of a carrier is typically abbreviated as AM-DSB/WC (amplitude modulation, double sideband/with carrier) and, when the carrier is suppressed or absent, as AM-DSB/SC (amplitude modulation, double-sideband/suppressed carrier). The corresponding single-sideband systems are abbreviated AM-SSB/WC and AM-SSB/SC.

Sections 8.1 through 8.4 are intended to provide an introduction to many of the basic concepts associated with sinusoidal amplitude modulation. There are many variations in details and implementation, and the reader is referred to the bibliography for an indication of the numerous excellent books that explore this topic further.

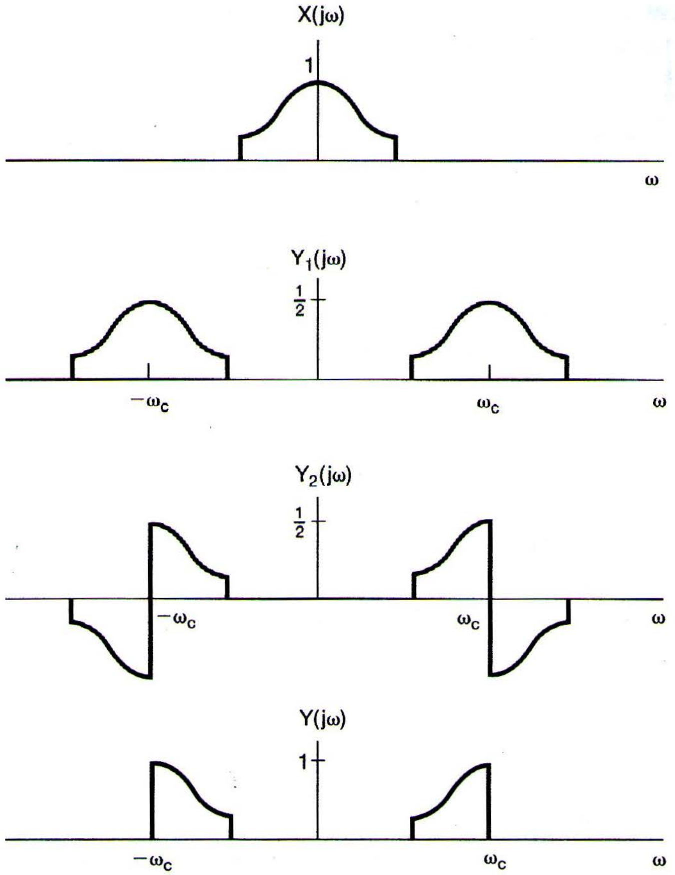


Figure 8.22 Spectra associated with the single-sideband system of Figure 8.21.

8.5 AMPLITUDE MODULATION WITH A PULSE-TRAIN CARRIER

8.5.1 Modulation of a Pulse-Train Carrier

In previous sections, we examined amplitude modulation with a sinusoidal carrier. Another important class of amplitude modulation techniques corresponds to the use of a carrier signal that is a pulse train, as illustrated in Figure 8.23; amplitude modulation of this type effectively corresponds to transmitting equally spaced time slices of $x(t)$. In general, we would not expect that an arbitrary signal could be recovered from such a set of time slices. However, our examination of the concept of sampling in Chapter 7 suggests that this should be possible if $x(t)$ is band limited and the pulse repetition frequency is high enough.

From Figure 8.23,

$$y(t) = x(t)c(t); \tag{8.22}$$

i.e., the modulated signal $y(t)$ is the product of $x(t)$ and the carrier $c(t)$. With $Y(j\omega)$, $X(j\omega)$, and $C(j\omega)$ representing the Fourier transforms of each of these signals, it follows from

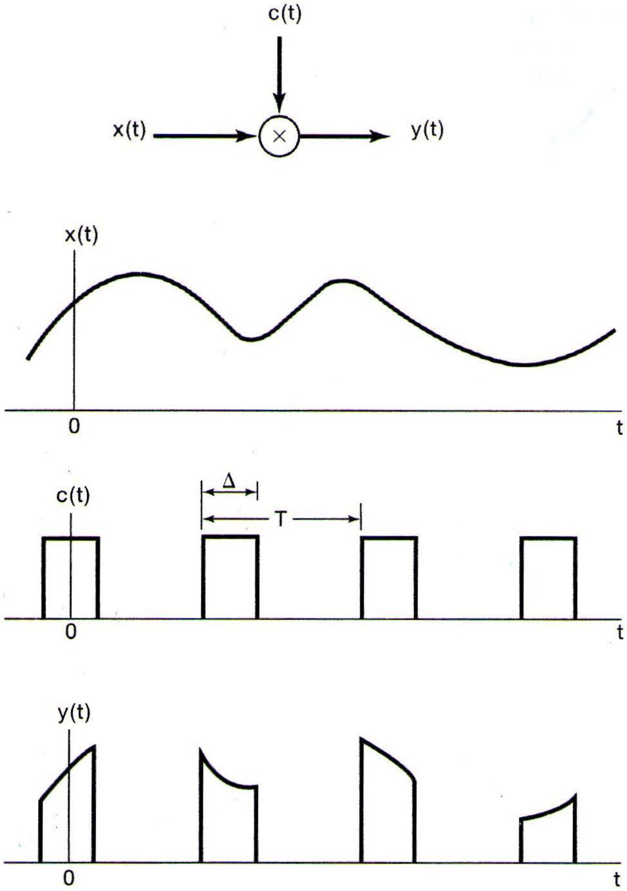


Figure 8.23 Amplitude modulation of a pulse train.

the multiplication property that

$$Y(j\omega) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} X(j\theta)C(j(\omega - \theta))d\theta. \tag{8.23}$$

Since $c(t)$ is periodic with period T, $C(j\omega)$ consists of impulses in frequency spaced by $2\pi/T$; that is,

$$C(j\omega) = 2\pi \sum_{k=-\infty}^{-\infty} a_k\delta(\omega - k\omega_c), \tag{8.24}$$

where $\omega_c = 2\pi/T$ and the coefficients a_k are the Fourier series coefficients of $c(t)$, which, from Example 3.5, are

$$a_k = \frac{\sin(k\omega_c\Delta/2)}{\pi k}. \tag{8.25}$$

The spectrum of $c(t)$ is shown in Figure 8.24(b). With the spectrum of $x(t)$ as illustrated in Figure 8.24(a), the resulting spectrum of the modulated signal $y(t)$ is shown in Figure 8.24(c). From eqs. (8.23) and (8.24), $Y(j\omega)$ is a sum of scaled and shifted replicas of $X(j\omega)$:

$$Y(j\omega) = \sum_{k=-\infty}^{+\infty} a_k X(j(\omega - k\omega_c)). \tag{8.26}$$

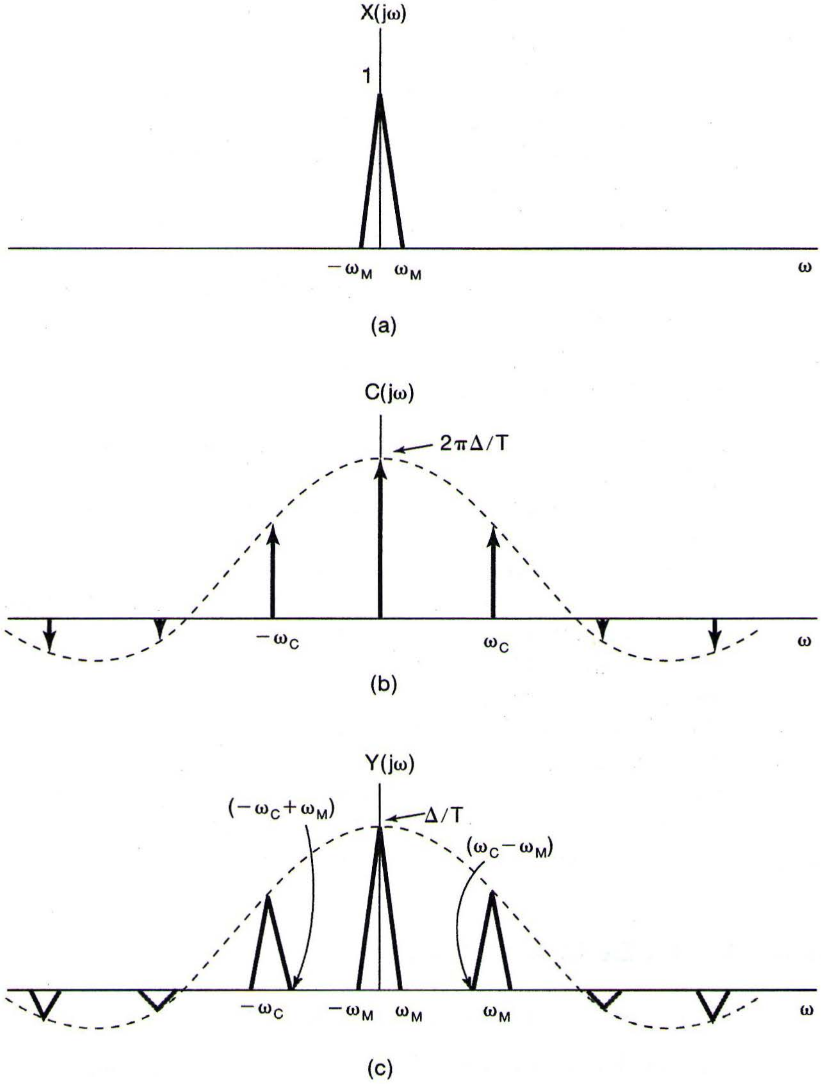


Figure 8.24 Spectra associated with amplitude modulation of a pulse train: (a) spectrum of a bandlimited-signal $x(t)$; (b) spectrum of the pulse carrier signal $c(t)$ in Figure 8.23; (c) spectrum of the modulated pulse train $y(t)$.

Comparing eq. (8.26) with eq. (7.6) and Figure 8.24 with Figure 7.3(c), we see that the spectrum of $y(t)$ is very similar in form to the spectrum resulting from sampling with a periodic impulse train, the only difference being the values of the Fourier coefficients of the pulse train. For the periodic impulse train used in Chapter 7, all of the Fourier coefficients are equal to $1/T$ in value, while for the pulse train $c(t)$ in Figure 8.23, the Fourier coefficients are given by eq. (8.25). Consequently, the replicas of $X(j\omega)$ do not overlap as long as $\omega_c > 2\omega_M$, which corresponds to the condition of the Nyquist sampling theorem. If this constraint is satisfied, then, as with impulse-train sampling, $x(t)$ can be recovered from $y(t)$ through the use of a lowpass filter with cutoff frequency greater than ω_M and less than $\omega_c - \omega_M$.

Note that the same conclusion holds for a wide variety of other pulselike carrier waveforms: If $c(t)$ is *any* periodic signal with Fourier transform as in eq. (8.24) for some set of Fourier coefficients a_k, then $Y(j\omega)$ is given by eq. (8.26). Then, as long as $\omega_c = 2\pi/T > 2\omega_M$, the replicas of $X(j\omega)$ do not overlap, allowing us to recover $x(t)$ by lowpass filtering, provided that the DC Fourier coefficient a_0 is nonzero. As shown in Problem 8.11, if a_0 is zero or unacceptably small, then, by using a bandpass filter to select one of the shifted replicas of $X(j\omega)$ with a larger value of a_k, we obtain a sinusoidal AM signal with a scaled version of $x(t)$ as the modulating signal. Using the demodulation methods described in Section 8.2, we can then recover $x(t)$.

8.5.2 Time-Division Multiplexing

Amplitude modulation with a pulse-train carrier is often used to transmit several signals over a single channel. As indicated in Figure 8.23, the modulated output signal $y(t)$ is nonzero only when the carrier signal $c(t)$ is on (i.e., is nonzero). During the intervals in which $c(t)$ is off, other similarly modulated signals can be transmitted. Two equivalent representations of this process are shown in Figure 8.25. In this technique for transmitting several signals over a single channel, each signal is in effect assigned a set of time slots of duration Δ that repeat every T seconds and that do not overlap with the slots assigned to other signals. The smaller the ratio Δ/T, the larger the number of signals that can be transmitted over the channel. This procedure is referred to as *time-division multiplexing* (TDM). Whereas frequency-division multiplexing, as discussed in Section 8.3, assigns different *frequency* intervals to individual signals, time-division multiplexing assigns different *time* intervals to individual signals. Demultiplexing the individual signals from the composite signal in Figure 8.25 is accomplished by time gating, to select the particular time slots associated with each individual signal.

8.6 PULSE-AMPLITUDE MODULATION

8.6.1 Pulse-Amplitude Modulated Signals

In Section 8.5 we described a modulation system in which a continuous-time signal $x(t)$ modulates a periodic pulse train, corresponding to transmitting time slices of $x(t)$ of duration Δ seconds every T seconds. As we saw both in that discussion and in our investigation of sampling in Chapter 7, our ability to recover $x(t)$ from these time slices depends not on their duration Δ, but rather on their frequency $2\pi/T$, which must exceed the Nyquist rate

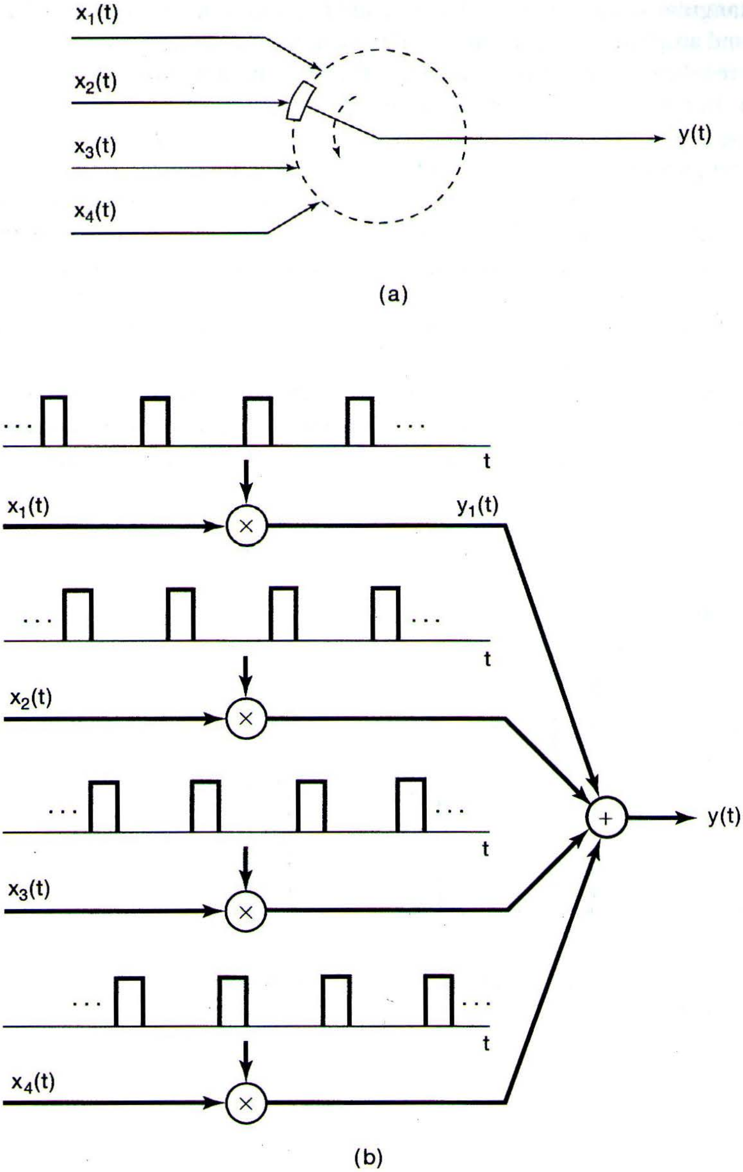


Figure 8.25 Time-division multiplexing.

in order to ensure an alias-free reconstruction of $x(t)$. That is, in principle, we need only transmit the samples $x(nT)$ of the signal $x(t)$.

In fact, in modern communication systems, sampled values of the information-bearing signal $x(t)$, rather than time slices are more typically transmitted. For practical reasons, there are limitations on the maximum amplitude that can be transmitted over a communication channel, so that transmitting impulse-sampled versions of $x(t)$ is not practical. Instead, the samples $x(nT)$ are used to modulate the amplitude of a sequence of pulses, resulting in what is referred to as a *pulse-amplitude modulation* (PAM) system.

The use of rectangular pulses corresponds to a sample-and-hold strategy in which pulses of duration Δ and amplitude proportional to the instantaneous sample values of $x(t)$ are transmitted. The resulting waveform for a single PAM channel of this type is illustrated in Figure 8.26. In the figure, the dotted curve represents the signal $x(t)$. As with the modulation scheme in Section 8.5, PAM signals can be time multiplexed. This is illustrated in Figure 8.27, which depicts the transmitted waveform with three time-multiplexed channels. The pulses associated with each channel are distinguished by shading, as well as by the channel number above each pulse. For a given pulse-repetition period T, as the pulse width decreases, more time-multiplexed channels can be transmitted over the same communication channel or medium. However, as the pulse width decreases, it is typically necessary to increase the amplitude of the transmitted pulses so that a reasonable amount of energy is transmitted in each pulse.

In addition to energy considerations, a number of other issues must be addressed in designing a PAM signal. In particular, as long as the sampling frequency exceeds the Nyquist rate, we know that $x(t)$ can be reconstructed exactly from its samples, and con-

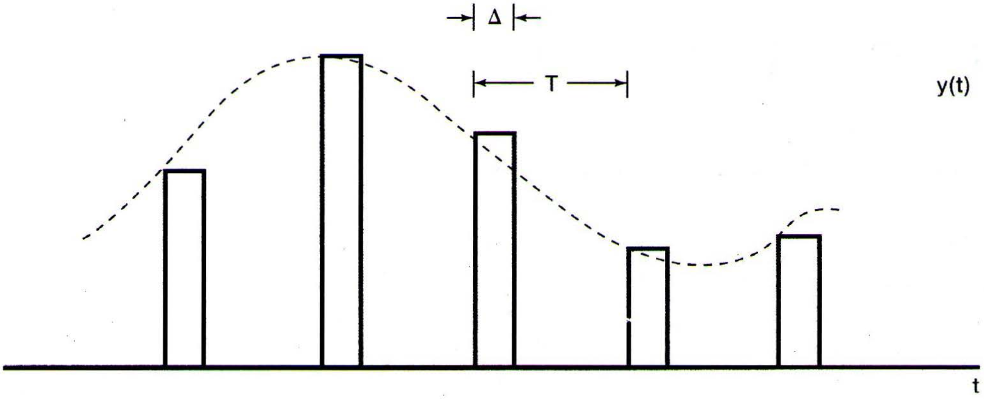


Figure 8.26 Transmitted waveform for a single PAM channel. The dotted curve represents the signal $x(t)$.

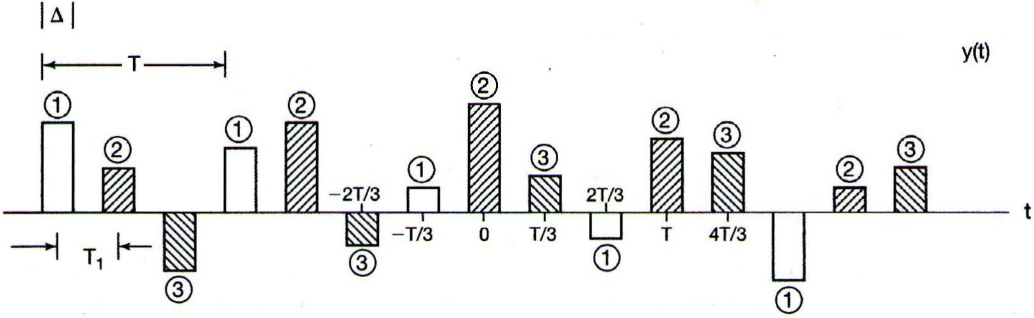


Figure 8.27 Transmitted waveform with three time-multiplexed PAM channels. The pulses associated with each channel are distinguished by shading, as well as by the channel number above each pulse. Here, the intersymbol spacing is $T_1 = T/3$.

sequently we can use these samples to modulate the amplitude of a sequence of pulses of any shape. The choice of pulse shape is dictated by considerations such as the frequency selectivity of the communication medium being used and the problem of intersymbol interference, which we discuss next.

8.6.2 Intersymbol Interference in PAM Systems

In the TDM pulse-amplitude modulation system just described, the receiver can, in principle, separate the channels by sampling the time-multiplexed waveform at appropriate times. For example, consider the time-multiplexed signal in Figure 8.27, which consists of pulse-amplitude-modulated versions of three signals $x_1(t)$, $x_2(t)$, and $x_3(t)$. If we sample $y(t)$ at appropriate times, corresponding, for example, to the midpoints of each pulse, we can separate the samples of the three signals. That is,

$$\begin{aligned} y(t) &= Ax_1(t), \qquad t = 0, \pm 3T_1, \pm 6T_1, \ldots, \\ y(t) &= Ax_2(t), \qquad t = T_1, T_1 \pm 3T_1, T_1 \pm 6T_1, \ldots, \\ y(t) &= Ax_3(t), \qquad t = 2T_1, 2T_1 \pm 3T_1, T_1 \pm 6T_1, \ldots, \end{aligned} \tag{8.27}$$

where T_1 is the intersymbol spacing, here equal to $T/3$, and where A is the appropriate proportionality constant. In other words, samples of $x_1(t)$, $x_2(t)$, and $x_3(t)$ can be obtained by appropriate sampling of the received time-multiplexed PAM signal.

The strategy indicated in the preceding paragraph assumes that the transmitted pulses remain distinct as they propagate over the communication channel. In transmission through any realistic channel, however, the pulses can be expected to be distorted through effects such as additive noise and filtering. Additive noise in the channel will, of course, introduce amplitude errors at the sampling times. Filtering due to the nonideal frequency response of a channel causes a smearing of the individual pulses that can cause the received pulses to overlap in time. This interference is illustrated in Figure 8.28 and is referred to as *intersymbol interference.*

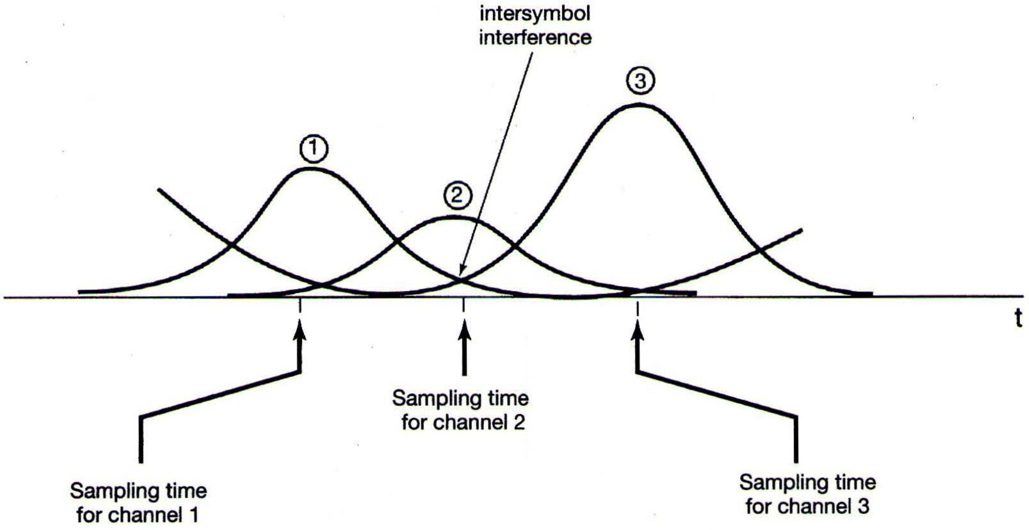


Figure 8.28 Intersymbol interference.

The smearing over time of the idealized pulses in Figure 8.27 can result from the bandwidth constraints of the channel or from phase dispersion caused by nonconstant group delay, as was discussed in Section 6.2.2. (See in particular, Example 6.1.) If the intersymbol interference is due only to the limited bandwidth of the channel, an approach is to use a pulse shape $p(t)$ that is itself band limited and therefore not affected (or only minimally affected) by the restricted bandwidth of the channel. In particular, if the channel has a frequency response $H(j\omega)$ that has no distortion over a specified frequency band (e.g., if $H(j\omega) = 1$ for $|\omega| < W$), then if the pulse that is used is band limited (i.e., if $P(j\omega) = 0$ for $|\omega| \geq W$), each PAM signal will be received without distortion. On the other hand, by using such a pulse, we no longer have pulses without overlap as in Figure 8.27. Nevertheless, intersymbol interference can be avoided in the time domain, even with a band-limited pulse, if the pulse shape is constrained to have zero-crossings at the other sampling times [so that eq. (8.27) continues to hold]. For example, consider the sinc pulse

$$p(t) = \frac{T_1 \sin(\pi t/T_1)}{\pi t}$$

and its corresponding spectrum displayed in Figure 8.29. Since the pulse is zero at integer multiples of the symbol spacing T_1, as indicated in Figure 8.30, there will be no intersymbol interference at these instants. That is, if we sample the received signal at $t = kT_1$, then the contributions to this sampled value from all of the other pulses, i.e., from $p(t - mT_1)$ for $m \neq k$, will be identically zero. Of course, avoiding interference from adjacent symbols

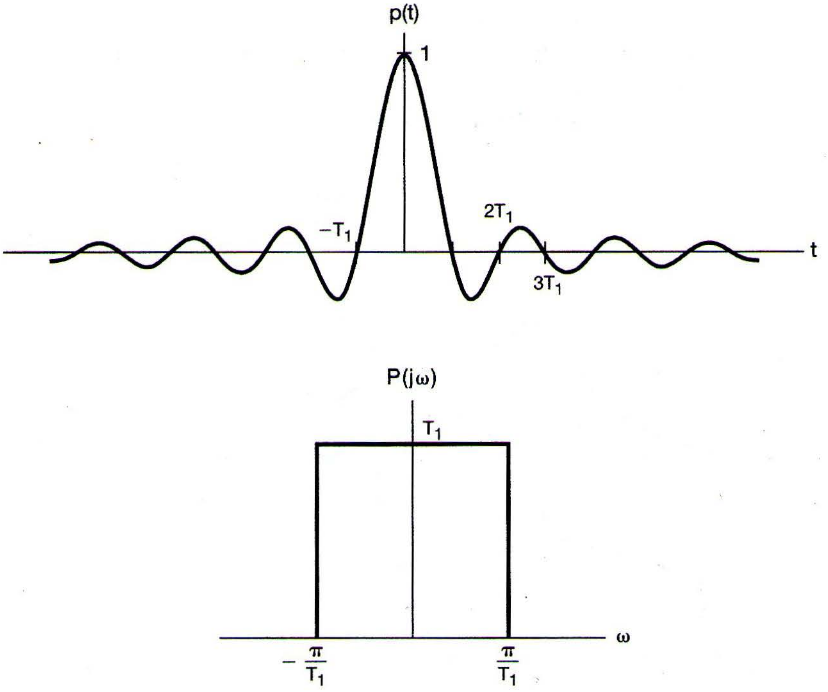


Figure 8.29 A sinc pulse and its corresponding spectrum.

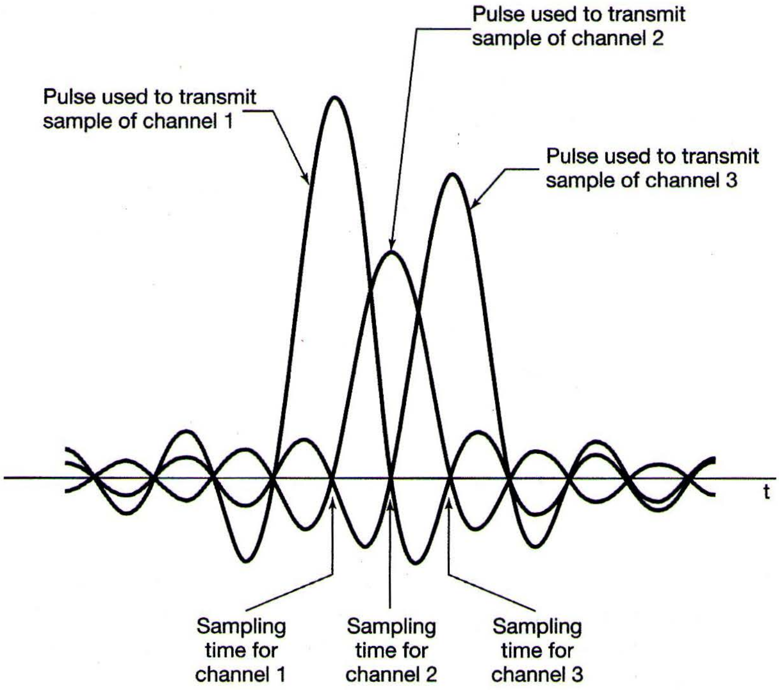

Pulse used to transmit sample of channel 1
Pulse used to transmit sample of channel 2
Pulse used to transmit sample of channel 3
t
Sampling time for channel 1
Sampling time for channel 2
Sampling time for channel 3

Figure 8.30 Absence of intersymbol interference when sinc pulses with correctly chosen zero-crossings are used.

requires high accuracy in the sampling times, so that sampling occurs at the zero-crossings of the adjacent symbols.

The sinc pulse is only one of many band-limited pulses with time-domain zero-crossings at $\pm T_1$, $\pm 2T_1$, etc. More generally, consider a pulse $p(t)$ with spectrum of the form

$$P(j\omega) = \begin{cases} 1 + P_1(j\omega), & |\omega| \le \frac{\pi}{T_1}, \\ P_1(j\omega), & \frac{\pi}{T_1} < |\omega| \le \frac{2\pi}{T_1} \\ 0, & \text{otherwise} \end{cases} \tag{8.28}$$

and with $P_1(j\omega)$ having odd symmetry around π/T_1, so that

$$P_1\left(-j\omega + j\frac{\pi}{T_1}\right) = -P_1\left(j\omega + j\frac{\pi}{T_1}\right) \quad 0 \le \omega \le \frac{\pi}{T_1}, \tag{8.29}$$

as illustrated in Figure 8.31. If $P_1(j\omega) = 0$, $p(t)$ is the sinc pulse itself. More generally, as explored in Problem 8.42, for any $P(j\omega)$ satisfying the conditions in eqs. (8.28) and (8.29), $p(t)$ will have zero-crossing at $\pm T_1, \pm 2T_1, \ldots$.

While signals satisfying eqs. (8.28) and (8.29) allow us to overcome the problem of limited channel bandwidth, other channel distortions may occur that require a different choice of pulse waveform or some additional processing of the received signal prior to the separation of the different TDM signals. In particular, if $|H(j\omega)|$ is not constant over the passband, there may be a need to perform *channel equalization*—i.e., filtering of

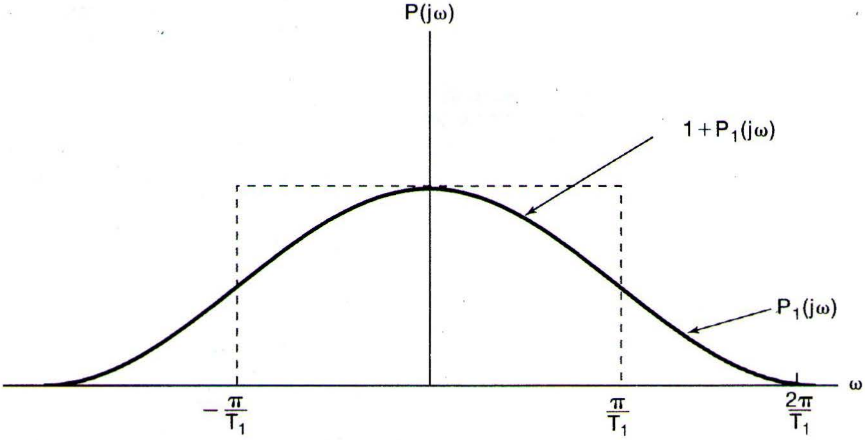


Figure 8.31 Odd symmetry around π/T_1 as defined in eq. (8.29).

the received signal to correct for the nonconstant channel gain. Also, if the channel has nonlinear phase, distortion can result that leads to intersymbol interference, unless compensating signal processing is performed. Problems 8.43 and 8.44 provide illustrations of these effects.

8.6.3 Digital Pulse-Amplitude and Pulse-Code Modulation

The PAM system described in the preceding subsections involves the use of a discrete set of samples to modulate a sequence of pulses. This set of samples can be thought of as a discrete-time signal $x[n]$, and in many applications $x[n]$ is in fact stored in or generated by a digital system. In such cases, the limited word length of a digital system implies that $x[n]$ can take on only a finite, quantized set of values, resulting in only a finite set of possible amplitudes for the modulated pulses.

In fact, in many cases this quantized form of digital PAM is reduced to a system using only a few—typically, only two—amplitude values. In particular, if each sample of $x[n]$ is represented as a binary number (i.e., a finite string of 0's and 1's), then a pulse with one of two possible values (one value corresponding to a 0 and one value to a 1) can be set for each binary digit, or *bit,* in the string. More generally, in order to protect against transmission errors or provide secure communication, the sequence of binary digits representing $x[n]$ might first be transformed or *encoded* into another sequence of 0's and 1's before transmission. For example, a very simple error detection mechanism is to transmit one additional modulated pulse for each sample of $x[n]$, representing a *parity* check. That is, this additional bit would be set to 1 if the binary representation of $x[n]$ has an odd number of 1's in it and to 0 if there is an even number of 1's. The receiver can then check the received parity bit against the other received bits in order to detect inconsistencies. More complex coding and error correction schemes can certainly be employed, and the design of codes with particular desirable properties is an important component of communication system design. For obvious reasons, a PAM system modulated by an encoded sequence of 0's and 1's is referred to as a *pulse-code modulation* (PCM) system.

8.7 SINUSOIDAL FREQUENCY MODULATION

In the preceding sections, we discussed a number of specific amplitude modulation systems in which the modulating signal was used to vary the amplitude of a sinusoidal or a pulse carrier. As we have seen, such systems are amenable to detailed analysis using the frequency-domain techniques we developed in preceding chapters. In another very important class of modulation techniques referred to as *frequency modulation (FM),* the modulating signal is used to control the frequency of a sinusoidal carrier. Modulation systems of this type have a number of advantages over amplitude modulation systems. As suggested by Figure 8.10, with sinusoidal amplitude modulation the peak amplitude of the envelope of the carrier is directly dependent on the amplitude of the modulating signal $x(t)$, which can have a large dynamic range—i.e., can vary significantly. With frequency modulation, the envelope of the carrier is constant. Consequently, an FM transmitter can always operate at peak power. In addition, in FM systems, amplitude variations introduced over a transmission channel due to additive disturbances or fading can, to a large extent, be eliminated at the receiver. For this reason, in public broadcasting and a variety of other contexts, FM reception is typically better than AM reception. On the other hand, as we will see, frequency modulation generally requires greater bandwidth than does sinusoidal amplitude modulation.

Frequency modulation systems are highly nonlinear and, consequently, are not as straightforward to analyze as are the amplitude modulation systems discussed in the preceding sections. However, the methods we have developed in earlier chapters do allow us to gain some understanding of the nature and operation of these systems.

We begin by introducing the general notion of *angle modulation.* Consider a sinusoidal carrier expressed in the form

$$c(t) = A\cos(\omega_c t + \theta_c) = A\cos\theta(t), \tag{8.30}$$

where $\theta(t) = \omega_c t + \theta_c$ and where ω_c is the frequency and θ_c the phase of the carrier. Angle modulation, in general, corresponds to using the modulating signal to change or vary the angle $\theta(t)$. One form that this sometimes takes is to use the modulating signal $x(t)$ to vary the phase θ_c so that the modulated signal takes the form

$$y(t) = A\cos[\omega_c t + \theta_c(t)], \tag{8.31}$$

where θ_c is now a function of time, specifically of the form

$$\theta_c(t) = \theta_0 + k_p x(t). \tag{8.32}$$

If $x(t)$ is, for example, constant, the phase of $y(t)$ will be constant and proportional to the amplitude of $x(t)$. Angle modulation of the form of eq. (8.31) is referred to as *phase modulation.* Another form of angle modulation corresponds to varying the *derivative* of the angle proportionally with the modulating signal; that is,

$$y(t) = A\cos\theta(t), \tag{8.33}$$

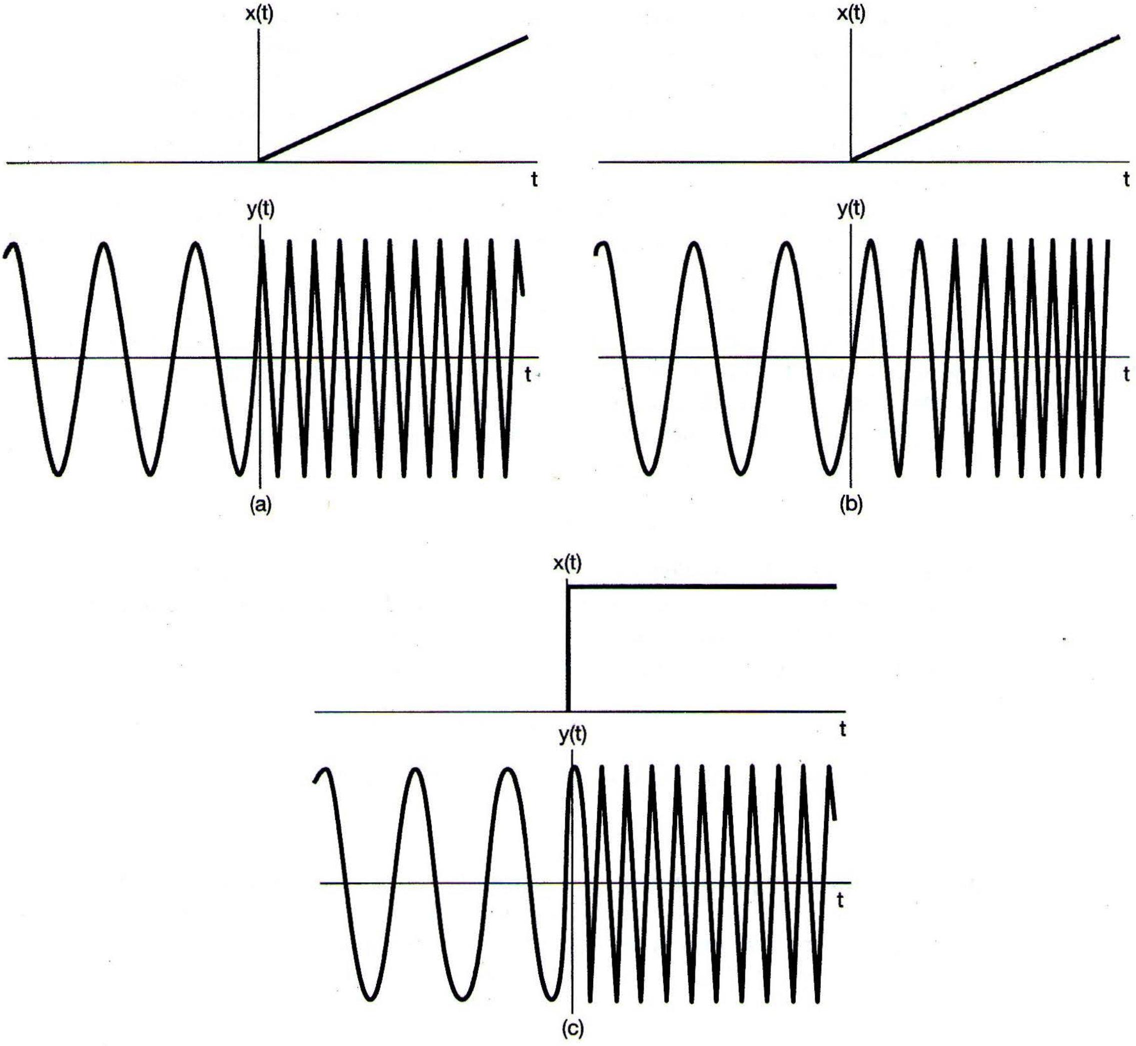
where

$$\frac{d\theta(t)}{dt} = \omega_c + k_f x(t). \tag{8.34}$$

For $x(t)$ constant, $y(t)$ is sinusoidal with a frequency that is offset from the carrier frequency ω_c by an amount proportional to the amplitude of $x(t)$. For that reason, angle modulation of the form of eqs. (8.33) and (8.34) is commonly referred to as *frequency modulation.*

Although phase modulation and frequency modulation are different forms of angle modulation, they can be easily related. From eqs. (8.31) and (8.32), for phase modulation,

$$\frac{d\theta(t)}{dt} = \omega_c + k_p \frac{dx(t)}{dt}, \tag{8.35}$$



(a)

(b)

(c)

Figure 8.32 Phase modulation, frequency modulation, and their relationship: (a) phase modulation with a ramp as the modulating signal; (b) frequency modulation with a ramp as the modulating signal; (c) frequency modulation with a step (the derivative of a ramp) as the modulating signal.

and thus, comparing eqs. (8.34) and (8.35), we see that phase modulating with $x(t)$ is identical to frequency modulating with the derivative of $x(t)$. Likewise, frequency modulating with $x(t)$ is identical to phase modulating with the integral of $x(t)$. An illustration of phase modulation and frequency modulation is shown in Figures 8.32(a) and (b). In both cases, the modulating signal is $x(t) = tu(t)$ (i.e., a ramp signal increasing linearly with time for $t > 0$). In Figure 8.32(c), an example of frequency modulation is shown with a step (the derivative of a ramp) as the modulating signal [i.e., $x(t) = u(t)$]. The correspondence between Figures 8.32(a) and (c) should be evident.

Frequency modulation with a step corresponds to the frequency of the sinusoidal carrier changing instantaneously from one value to another when $x(t)$ changes value at $t = 0$, much as the frequency of a sinusoidal oscillator changes when the frequency setting is switched instantaneously. When the frequency modulation is a ramp, as in Figure 8.32(b), the frequency changes linearly with time. This notion of a time-varying frequency is often best expressed in terms of the concept of *instantaneous frequency.* For

$$y(t) = A\cos\theta(t), \tag{8.36}$$

the instantaneous frequency of the sinusoid is defined as

$$\omega_i(t) = \frac{d\theta(t)}{dt}. \tag{8.37}$$

Thus, for $y(t)$ truly sinusoidal [i.e., $\theta(t) = (\omega_c t + \theta_0)$], the instantaneous frequency is ω_c, as we would expect. For phase modulation as expressed in eqs. (8.31) and (8.32), the instantaneous frequency is $\omega_c + k_p(dx(t)/dt)$, and for frequency modulation as expressed in eqs. (8.33) and (8.34), the instantaneous frequency is $\omega_c + k_f x(t)$.

Since frequency modulation and phase modulation are easily related, we will phrase the remaining discussion in terms of frequency modulation alone. To gain some insight into how the spectrum of the frequency-modulated signal is affected by the modulating signal $x(t)$, it is useful to consider two cases in which the modulating signal is sufficiently simple so that some of the essential properties of frequency modulation become evident.

8.7.1 Narrowband Frequency Modulation

Consider the case of frequency modulation with

$$x(t) = A\cos\omega_m t. \tag{8.38}$$

From eqs. (8.34) and (8.37), the instantaneous frequency is

$$\omega_i(t) = \omega_c + k_f A\cos\omega_m t, \tag{8.39}$$

which varies sinusoidally between $\omega_c + k_f A$ and $\omega_c - k_f A$. With

$$\Delta\omega = k_f A,$$

we have

$$\omega_i(t) = \omega_c + \Delta\omega\cos\omega_m t,$$

and

$$\begin{aligned} y(t) &= \cos[\omega_c t + \int x(t)dt] \\ &= \cos\left(\omega_c t + \frac{\Delta\omega}{\omega_m}\sin\omega_m t + \theta_0\right), \end{aligned} \tag{8.40}$$

where θ_0 is a constant of integration. For convenience we will choose $\theta_0 = 0$, so that

$$y(t) = \cos\left[\omega_c t + \frac{\Delta\omega}{\omega_m}\sin\omega_m t\right]. \tag{8.41}$$

The factor $\Delta\omega/\omega_m$, which we denote by m, is defined as the *modulation index* for frequency modulation. The properties of FM systems tend to be different, depending on whether the modulation index m is small or large. The case in which m is small is referred to as *narrowband FM*. In general, we can rewrite eq. (8.41) as

$$y(t) = \cos(\omega_c t + m\sin\omega_m t) \tag{8.42}$$

or

$$y(t) = \cos\omega_c t\cos(m\sin\omega_m t) - \sin\omega_c t\sin(m\sin\omega_m t). \tag{8.43}$$

When m is sufficiently small ($\ll \pi/2$), we can make the approximations

$$\cos(m\sin\omega_m t) \simeq 1, \tag{8.44}$$

$$\sin(m\sin\omega_m t) \simeq m\sin\omega_m t, \tag{8.45}$$

so that eq. (8.42) becomes

$$y(t) \simeq \cos\omega_c t - m(\sin\omega_m t)(\sin\omega_c t). \tag{8.46}$$

The spectrum of $y(t)$ based on this approximation is shown in Figure 8.33. We note that it has a similarity to AM-DSB/WC in that the carrier frequency is present in the spectrum and there are sidebands representing the spectrum of the modulating signal in eq. (8.38). However, in AM-DSB/WC the additional carrier injected is in phase with the modulated carrier, whereas, as we see in eq. (8.46) for the case of the narrowband FM, the carrier signal has a phase difference of $\pi/2$ in relation to the amplitude-modulated carrier. The waveforms corresponding to AM-DSB/WC and FM are also very different. Figure 8.34(a) illustrates the narrowband FM waveform corresponding to eq. (8.46). For comparison, Figure 8.34(b) shows the AM-DSB/WC signal

$$y_2(t) = \cos\omega_c t + m(\cos\omega_m t)(\cos\omega_c t). \tag{8.47}$$

For the narrowband FM signal of eq. (8.46), the bandwidth of the sidebands is equal to the bandwidth of the modulating signal, and in particular, although the approximation in the equation is based on assuming that $m \ll \pi/2$, the bandwidth of the sidebands is otherwise independent of the modulation index m (i.e., it depends only on the bandwidth of the modulating signal, not on its amplitude). A similar statement applies for narrowband FM with a more general modulating signal.

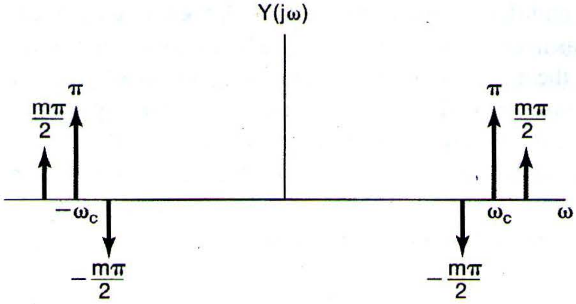


Figure 8.33 Approximate spectrum for narrowband FM.

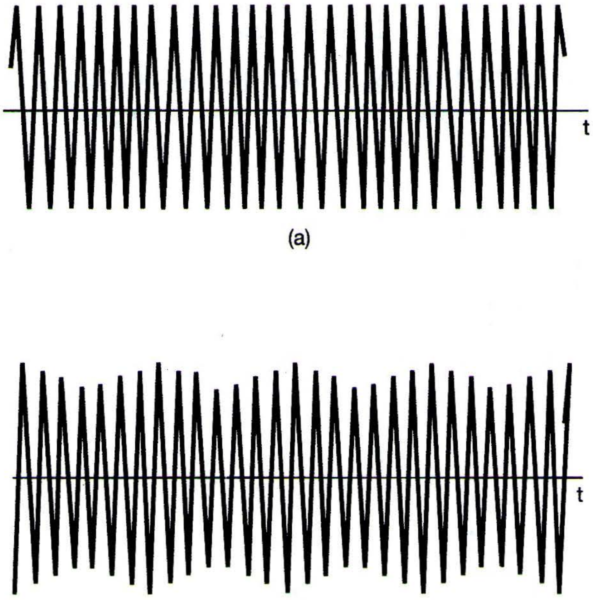


Figure 8.34 Comparison of narrowband FM and AM-DSB/WC: (a) narrowband FM; (b) AM-DSB/WC.

8.7.2 Wideband Frequency Modulation

When m is large, the approximation leading to eq. (8.46) no longer applies, and the spectrum of $y(t)$ depends on both the amplitude and the spectrum of the modulating signal $x(t)$. With $y(t)$ expressed in the form of eq. (8.43), we note that the terms $\cos[m \sin \omega_m t]$ and $\sin[m \sin \omega_m t]$ are periodic signals with fundamental frequency ω_m. Thus, the Fourier transform of each of these signals is an impulse train with impulses at integer multiples of ω_m and amplitudes proportional to the Fourier series coefficients. The coefficients for these two periodic signals involve a class of functions referred to as Bessel functions of the first kind. The first term in eq. (8.43) corresponds to a sinusoidal carrier of the form $\cos \omega_c t$ amplitude modulated by the periodic signal $\cos[m \sin \omega_m t]$ and the second term to a sinusoidal carrier $\sin \omega_c t$ amplitude modulated by the periodic signal $\sin[m \sin \omega_m t]$. Multiplication by the carrier signals has the effect in the frequency domain of translating the spectrum of eq. (8.43) to the carrier frequency, so that it is centered at plus and minus

ω_c. In Figures 8.35(a) and (b) we illustrate, for $\omega > 0$, the the magnitude of the spectra of the two individual terms in eq. (8.43), and in Figure 8.35(c) the magnitude of the combined spectrum representing the modulated signal $y(t)$. The spectrum of $y(t)$ consists of impulses at frequencies $\pm\omega_c + n\omega_m$, $n = 0, \pm 1, \pm 2, \ldots$, and is not, strictly speaking, band limited around $\pm\omega_c$. However, the behavior of the Fourier series coefficients of $\cos[m \sin \omega_m t]$ and $\sin[m \sin \omega_m t]$ are such that the amplitude of the nth harmonic for $|n| > m$ can be considered negligible, and thus, the total bandwidth B of each sideband centered around $+\omega_c$ and $-\omega_c$ is effectively limited to $2m\omega_m$. That is,

$$B \simeq 2m\omega_m, \tag{8.48}$$

or, since $m = k_f A/\omega_m = \Delta\omega/\omega_m$,

$$B \simeq 2k_f A = 2\Delta\omega. \tag{8.49}$$

Comparing eqs. (8.39) and (8.49), we note that the effective bandwidth of each sideband is equal to the total excursion of the instantaneous frequency around the carrier

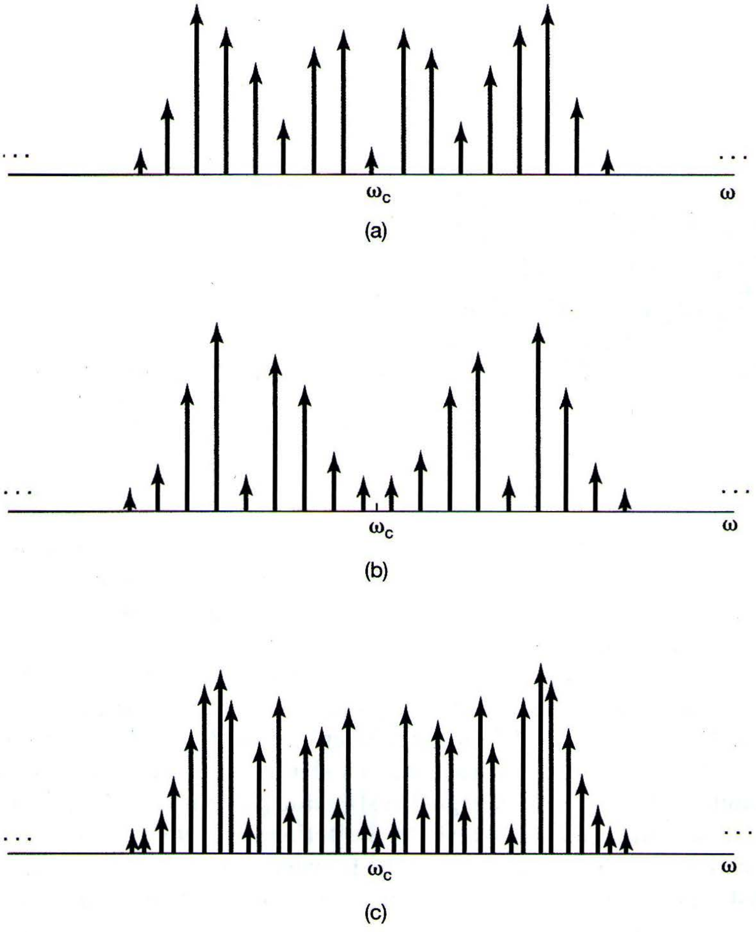


Figure 8.35 Magnitude of spectrum of wideband frequency modulation with $m = 12$: (a) magnitude of spectrum of $\cos \omega_c t \cos[m \sin \omega_m t]$; (b) magnitude of spectrum of $\sin \omega_c t \sin[m \sin \omega_m t]$; (c) combined spectral magnitude of $\cos[\omega_c t + m \sin \omega_m t]$.

frequency. Therefore, for wideband FM, since we assume that m is large, the bandwidth of the modulated signal is much larger than the bandwidth of the modulating signal, and in contrast to the narrowband case, the bandwidth of the transmitted signal in wideband FM is directly proportional to amplitude A of the modulating signal and the gain factor k_f.

8.7.3 Periodic Square-Wave Modulating Signal

Another example that lends insight into the properties of frequency modulation is that of a modulating signal which is a periodic square wave. Referring to eq. (8.39), let $k_f = 1$ so that $\Delta\omega = A$, and let $x(t)$ be given by Figure 8.36. The modulated signal $y(t)$ is illustrated in Figure 8.37. The instantaneous frequency is $\omega_c + \Delta\omega$ when $x(t)$ is positive and $\omega_c - \Delta\omega$ when $x(t)$ is negative. Thus, $y(t)$ can also be written as

$$y(t) = r(t)\cos[(\omega_c + \Delta\omega)t] + r\left(t - \frac{T}{2}\right)\cos[(\omega_c - \Delta\omega)t], \tag{8.50}$$

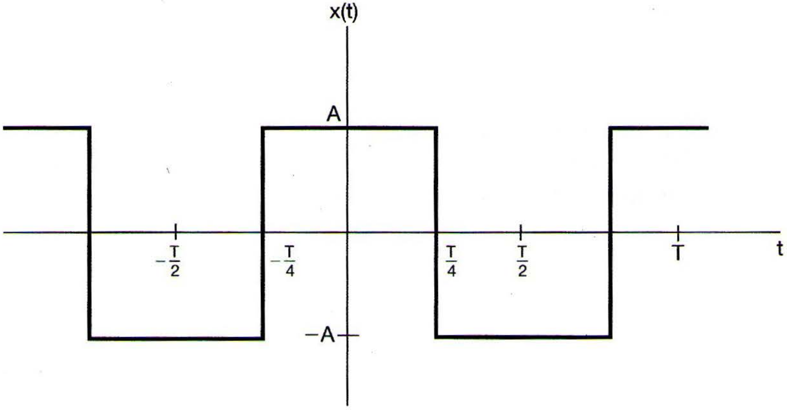


Figure 8.36 Symmetric periodic square wave.

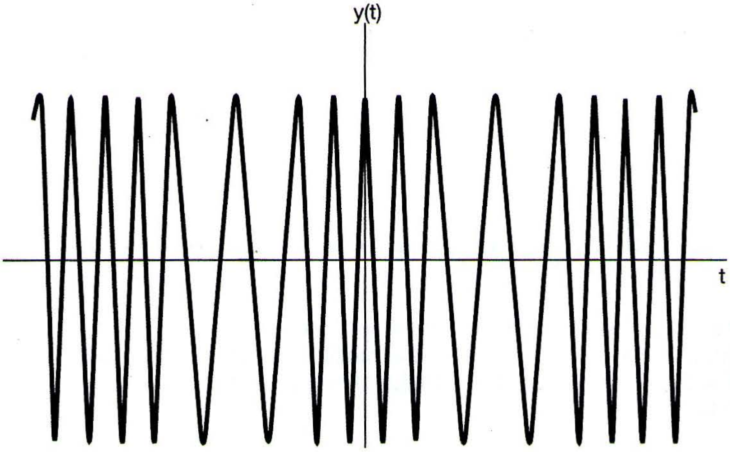


Figure 8.37 Frequency modulation with a periodic square-wave modulating signal.

where $r(t)$ is the symmetric square wave shown in Figure 8.38. Thus, for this particular modulating signal, we are able to recast the problem of determining the spectrum of the FM signal $y(t)$ as the determination of the spectrum of the sum of the two AM signals in eq. (8.50). Specifically,

$$\begin{aligned} Y(j\omega) = {} & \frac{1}{2}[R(j\omega + j\omega_c + j\Delta\omega) + R(j\omega - j\omega_c - j\Delta\omega)] \\ & + \frac{1}{2}[R_T(j\omega + j\omega_c - j\Delta\omega) + R_T(j\omega - j\omega_c + j\Delta\omega)], \end{aligned} \tag{8.51}$$

where $R(j\omega)$ is the Fourier transform of the periodic square wave $r(t)$ in Figure 8.38 and $R_T(j\omega)$ is the Fourier transform of $r(t - T/2)$. From Example 4.6, with $T = 4T_1$,

$$R(j\omega) = \sum_{k=-\infty}^{\infty} \frac{2}{2k+1}(-1)^k \delta\left[\omega - \frac{2\pi(2k+1)}{T}\right] + \pi\delta(\omega) \tag{8.52}$$

and

$$R_T(j\omega) = R(j\omega)e^{-j\omega T/2}. \tag{8.53}$$

The magnitude of the spectrum of $Y(j\omega)$ is illustrated in Figure 8.39. As with wideband FM, the spectrum has the general appearance of two sidebands, centered around $\omega_c \pm \Delta\omega$, that decay for $\omega < \omega_c - \Delta\omega$ and $\omega > \omega_c + \Delta\omega$.

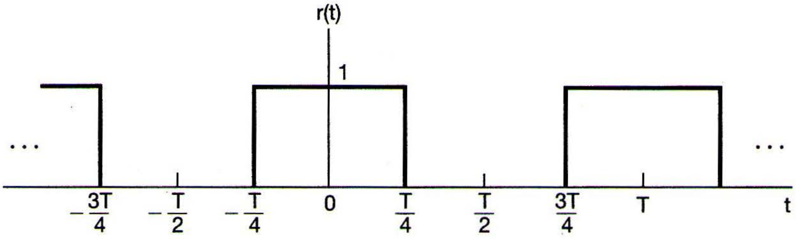


Figure 8.38 Symmetric square wave $r(t)$ in eq. (8.50).

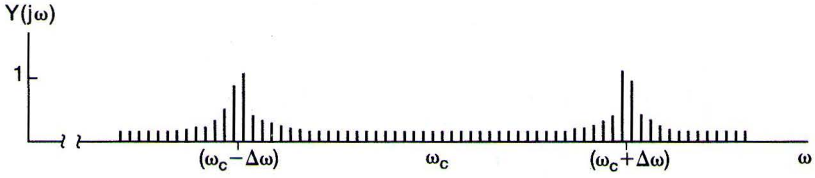


Figure 8.39 Magnitude of the spectrum for $\omega > 0$ corresponding to frequency modulation with a periodic square-wave modulating signal. Each of the vertical lines in the figure represents an impulse of area proportional to the height of the line.

Systems for the demodulation of FM signals typically are of two types. One type of demodulation system corresponds to converting the FM signal to an AM signal through differentiation, while demodulation systems of the second type directly track the phase or frequency of the modulated signal. The foregoing discussion provides only a brief introduction to the characteristics of frequency modulation, and we have again seen how the basic techniques developed in the earlier chapters can be exploited to analyze and gain an insight into an important class of systems.

8.8 DISCRETE-TIME MODULATION

8.8.1 Discrete-Time Sinusoidal Amplitude Modulation

A discrete-time amplitude modulation system is depicted in Figure 8.40, in which $c[n]$ is the carrier and $x[n]$ the modulating signal. The basis for our analysis of continuous-time amplitude modulation was the multiplication property for Fourier transforms—specifically, the fact that multiplication in the time domain corresponds to convolution in the frequency domain. As we discussed in Section 5.5, there is a corresponding property for discrete-time signals which we can use to analyze discrete-time amplitude modulation. Specifically, consider

$$y[n] = x[n]c[n].$$

With $X(e^{j\omega})$, $Y(e^{j\omega})$, and $C(e^{j\omega})$ denoting the Fourier transforms of $x[n]$, $y[n]$, and $c[n]$, respectively, $Y(e^{j\omega})$ is proportional to the periodic convolution of $X(e^{j\omega})$ and $C(e^{j\omega})$; that is,

$$Y(e^{j\omega}) = \frac{1}{2\pi}\int_{2\pi} X(e^{j\theta})C(e^{j(\omega-\theta)})d\theta. \tag{8.54}$$

Since $X(e^{j\omega})$ and $C(e^{j\omega})$ are periodic with a period of 2π, the integration can be performed over any frequency interval of length 2π.

Let us first consider sinusoidal amplitude modulation with a complex exponential carrier, so that

$$c[n] = e^{j\omega_c n}. \tag{8.55}$$

As we saw in Section 5.2, the Fourier transform of $c[n]$ is a periodic impulse train; that is,

$$C(e^{j\omega}) = \sum_{k=-\infty}^{+\infty} 2\pi\delta(\omega - \omega_c + k2\pi), \tag{8.56}$$

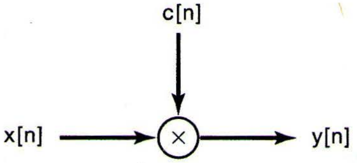


Figure 8.40 Discrete-time amplitude modulation.

which is sketched in Figure 8.41(b). With $X(e^{j\omega})$ as illustrated in Figure 8.41(a), the spectrum of the modulated signal is that shown in Figure 8.41(c). In particular, we note that $Y(e^{j\omega}) = X(e^{j(\omega-\omega_c)})$. This is the discrete-time counterpart to Figure 8.1, and here again, with $x[n]$ real, the modulated signal will be complex. Demodulation is accomplished by multiplying by $e^{-j\omega_c n}$ to translate the spectrum back to its original location on the frequency axis, so that

$$x[n] = y[n]e^{-j\omega_c n}. \tag{8.57}$$

As explored in Problem 6.43, if $\omega_c = \pi$ so that $c[n] = (-1)^n$, the result of modulation in the time domain is to change the algebraic sign of $x[n]$ for odd values of n, while in the frequency domain the consequence is the interchanging of high and low frequencies. Problem 6.44 explores the use of this type of modulation in utilizing a lowpass filter to achieve highpass filtering and vice versa.

As an alternative to a complex exponential carrier, we can use a sinusoidal carrier, in which case, with $x[n]$ real, the modulated signal $y[n]$ will also be real. With $c[n] = \cos\omega_c n$, the spectrum of the carrier consists of periodically repeated pairs of impulses at

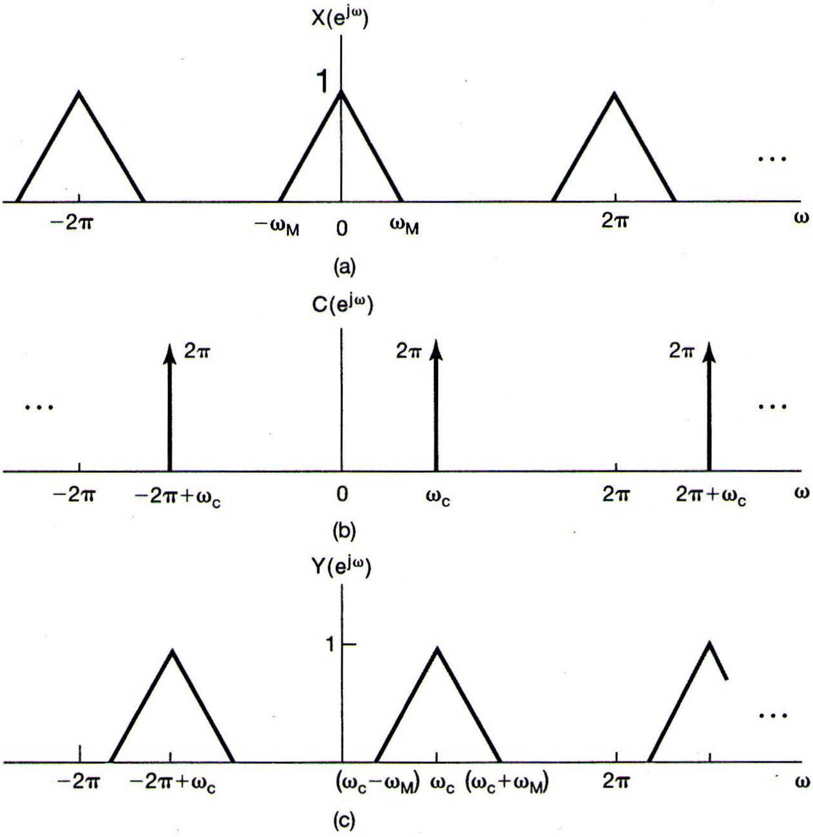


Figure 8.41 (a) Spectrum of $x[n]$; (b) spectrum of $c[n] = e^{j\omega_c n}$; (c) spectrum of $y[n] = x[n]c[n]$.

$\omega = \pm\omega_c + k2\pi$, as illustrated in Figure 8.42(b). With $X(e^{j\omega})$ as shown in Figure 8.42(a), the resulting spectrum for the modulated signal is shown in Figure 8.42(c) and corresponds to replicating $X(e^{j\omega})$ at the frequencies $\omega = \pm\omega_c + k2\pi$. In order that the individual replications of $X(e^{j\omega})$ do not overlap, we require that

$$\omega_c > \omega_M \tag{8.58}$$

and

$$2\pi - \omega_c - \omega_M > \omega_c + \omega_M$$

or, equivalently,

$$\omega_c < \pi - \omega_M. \tag{8.59}$$

The first condition is identical to that in Section 8.2 for continous-time sinusoidal amplitude modulation, while the second results from the inherent periodicity of discrete-time

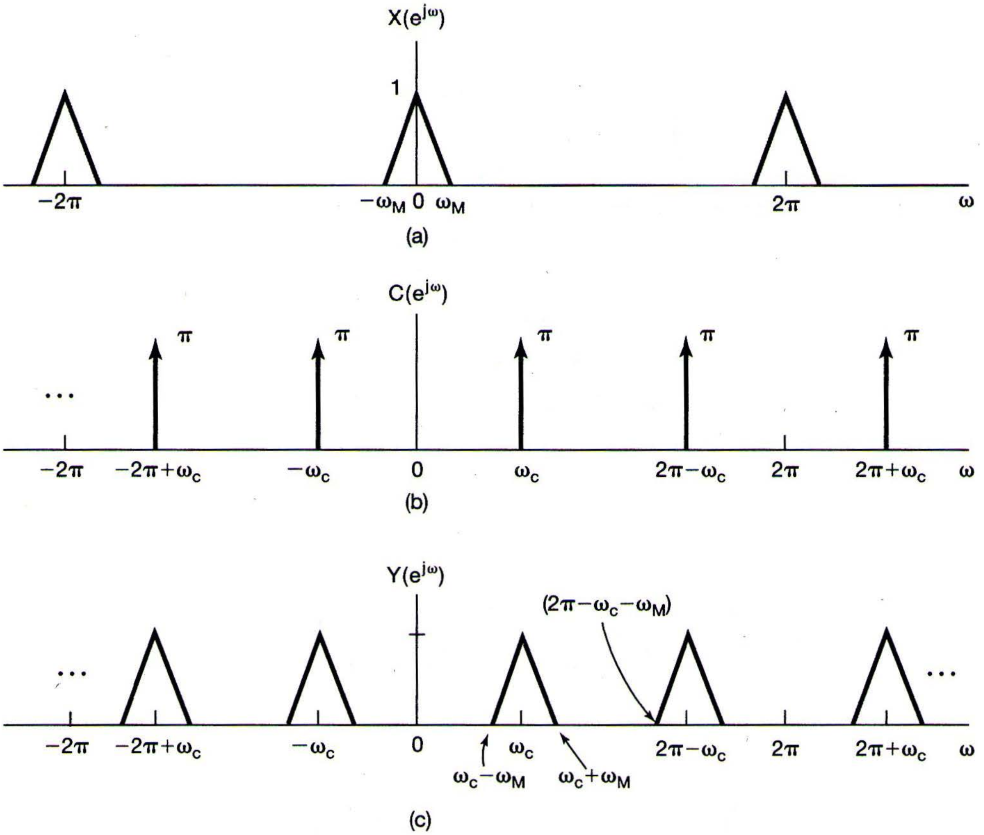


Figure 8.42 Spectra associated with discrete-time modulation using a sinusoidal carrier: (a) spectrum of a bandlimited-signal $x[n]$; (b) spectrum of a sinusoidal carrier signal $c[n] = \cos\omega_c n$; (c) spectrum of the modulated signal $y[n] = x[n]c[n]$.

spectra. Combining eqs. (8.58) and (8.59), we see that for amplitude modulation with a sinusoidal carrier, we must restrict ω_c so that

$$\omega_M < \omega_c < \pi - \omega_M. \tag{8.60}$$

Demodulation can be accomplished in a manner similar to that employed in continuous time. As illustrated in Figure 8.43, multiplication of $y[n]$ with the same carrier used in the modulator results in several replications of the spectrum of the original signal, one of which is centered about $\omega = 0$. By lowpass filtering to eliminate the unwanted replications of $X(e^{j\omega})$, the demodulated signal is obtained.

As should be evident from the foregoing discussion, analysis of discrete-time amplitude modulation proceeds in a manner similar to that of continuous-time amplitude modulation, with only slight differences. For example, as explored in Problem 8.47, in the synchronous modulation and demodulation system, the effect of a phase difference or a frequency difference between the sinusoidal carriers in the modulator and demodulator is identical in both discrete and continuous time. In addition, just as in continuous time, we can use discrete-time sinusoidal AM as the basis for frequency-division multiplexing in

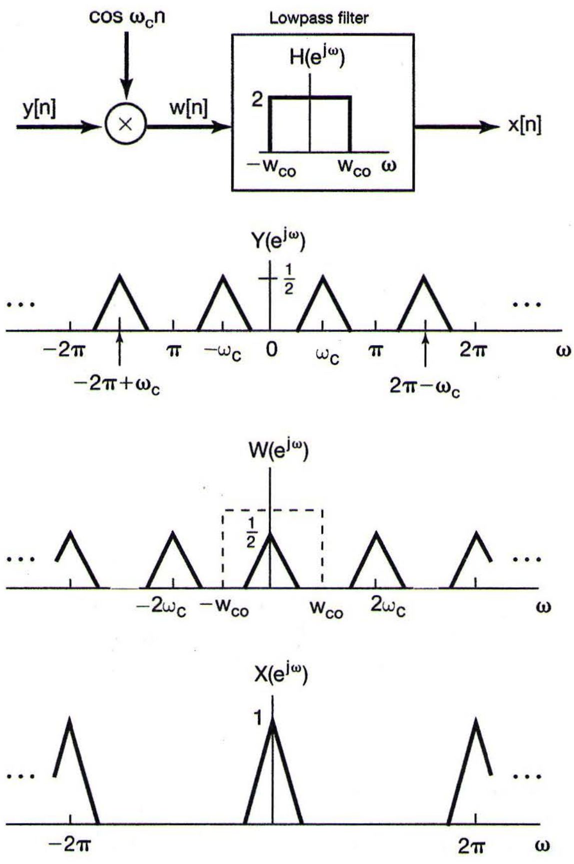


Figure 8.43 System and associated spectra for discrete-time synchronous demodulation.

discrete time. Furthermore, as explored in Problem 8.48, we can also consider using a discrete-time signal to modulate a pulse train, leading to time-division multiplexing of discrete-time signals.

The implementation of discrete-time multiplexing systems provides an excellent example of the flexibility of discrete-time processing in general and the importance of the operation of upsampling (see Section 7.5.2) in particular. Consider a discrete-time FDM system with M sequences that we wish to frequency-division multiplex. With M channels, it is required that the spectral energy for each input channel $x_i[n]$ be band limited; that is,

$$X_i(\omega) = 0, \qquad \frac{\pi}{M} < |\omega| < \pi. \tag{8.61}$$

If the sequences originally occupied the entire frequency band corresponding, for example, to having sampled a set of continuous-time signals at the Nyquist rate, then they would first have to be converted to a higher sampling rate (i.e., upsampled) before frequency-division multiplexing. This idea is explored further in Problem 8.33.

8.8.2 Discrete-Time Transmodulation

One context in which discrete-time modulation is widely used, together with the operations of decimation, upsampling, and interpolation introduced in Chapter 7, is digital communication systems. Typically, in such systems continuous-time signals are transmitted over communication channels in the form of discrete-time signals, obtained by sampling. The continous-time signals are often in the form of time-division-multiplexed (TDM) or frequency-division-multiplexed (FDM) signals. The signals are then converted to discrete-time sequences whose values are represented digitally, for storage or long-distance transmission. In some systems, because of different constraints or requirements at the transmitting end and the receiving end, or because sets of signals that have been individually multiplexed by different methods are then multiplexed together, there is often the requirement for converting from sequences representing TDM signals to sequences representing FDM signals or vice versa. This conversion from one modulation or multiplexing scheme to another is referred to as *transmodulation* or *transmultiplexing*. In the context of digital communication systems, one obvious way of implementing transmultiplexing is to convert back to continuous-time signals, demultiplex and demodulate, and then modulate and multiplex as required. However, if the new signal is then to be converted back to a discrete-time signal, it is clearly more efficient for the entire process to be carried out directly in the discrete-time domain. Figure 8.44 shows, in block diagram form, the steps involved in converting a discrete-time TDM signal to a discrete-time FDM signal. Note that, after demultiplexing the TDM signal, each channel must be upsampled in preparation for frequency-division multiplexing.

8.9 SUMMARY

In this chapter, we have examined a number of the basic concepts associated with communication systems. In particular, we have examined the concept of modulation, in which a signal we wish to communicate is used to modulate a second signal referred to as the

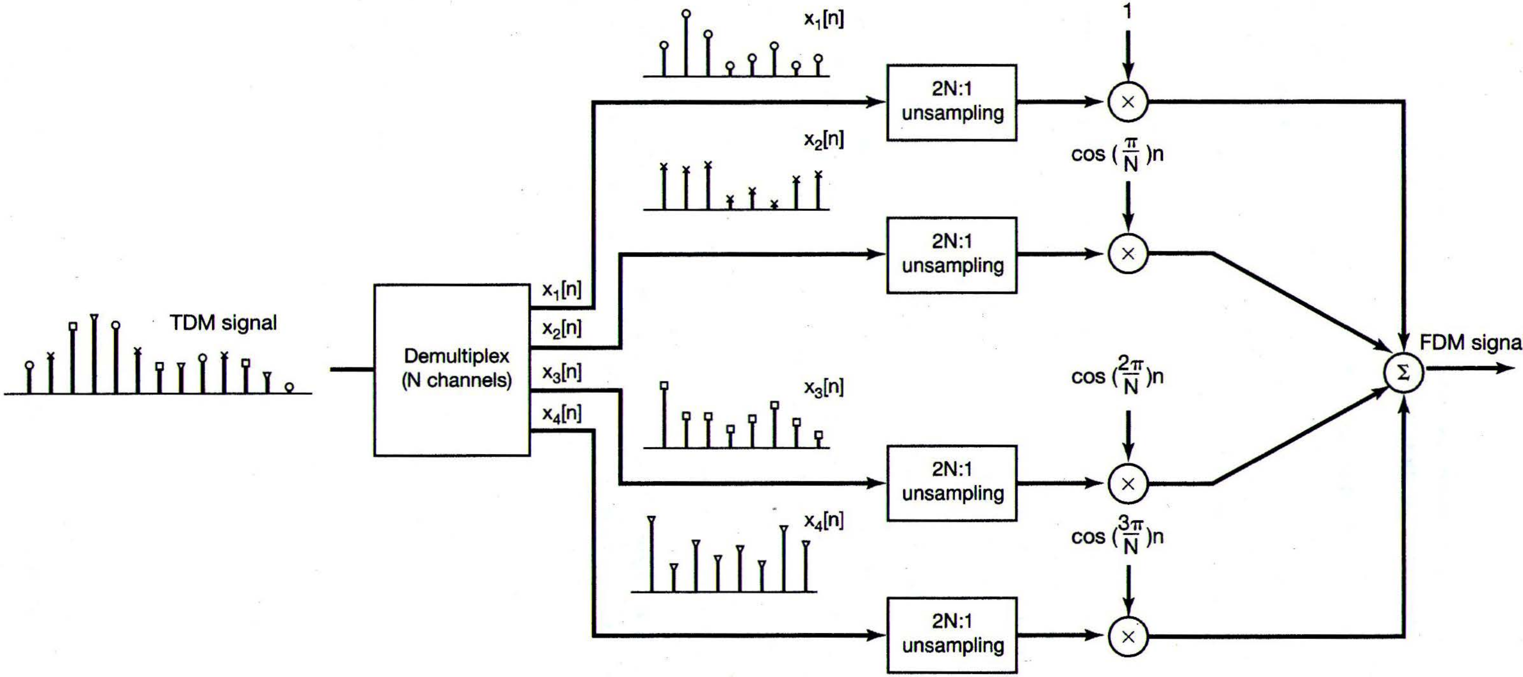

TDM signal
Demultiplex
(N channels)
x1[n]
x2[n]
x3[n]
x4[n]
2N:1
unsampling
1
cos (π/N)n
cos (2π/N)n
cos (3π/N)n
×
Σ
FDM signal

Figure 8.44 Block diagram for TDM-to-FDM transmultiplexing.

carrier, and we have looked in detail at the concept of amplitude modulation. The properties of amplitude modulation are most easily interpreted in the frequency domain through the multiplication property of the Fourier transform. Amplitude modulation with a complex exponential or sinusoidal carrier is typically used to shift the spectrum of a signal in frequency and is applied, for example, in communication systems to place the spectrum in a frequency range suitable for transmission and to permit frequency-division multiplexing. Variations of sinusoidal amplitude modulation, such as the insertion of a carrier signal for asynchronous systems and single- and double-sideband systems, were discussed.

We also examined several other forms of modulation-based communication. In this regard, we briefly introduced the concepts of frequency and phase modulation. Although these forms of modulation are more difficult to analyze in detail, it is possible to gain significant insight into their characteristics through the frequency domain.

We further examined in some detail amplitude modulation of a pulsed signal, which led us to the concepts of time-division multiplexing and pulse-amplitude modulation, in which successive samples of a discrete-time signal are used to modulate the amplitude of a sequence of pulses. This led in turn to an examination of discrete-time modulation and digital communication, in which the flexibility of discrete-time processing facilitates the design and implementation of more sophisticated communication systems involving concepts such as pulse-code modulation and transmodulation.

Chapter 8 Problems

The first section of problems belongs to the basic category, and the answers are provided in the back of the book. The remaining two sections contain problems belonging to the basic and advanced categories, respectively.

BASIC PROBLEMS WITH ANSWERS

8.1. Let $x(t)$ be a signal for which $X(j\omega) = 0$ when $|\omega| > \omega_M$. Another signal $y(t)$ is specified as having the Fourier transform $Y(j\omega) = 2X(j(\omega - \omega_c))$. Determine a signal $m(t)$ such that

$$x(t) = y(t)m(t).$$

8.2. Let $x(t)$ be a real-valued signal for which $X(j\omega) = 0$ when $|\omega| > 1{,}000\pi$. Supposing that $y(t) = e^{j\omega_c t}x(t)$, answer the following questions:

(a) What constraint should be placed on ω_c to ensure that $x(t)$ is recoverable from $y(t)$?

(b) What constraint should be placed on ω_c to ensure that $x(t)$ is recoverable from $\mathcal{R}e\{y(t)\}$?

8.3. Let $x(t)$ be a real-valued signal for which $X(j\omega) = 0$ when $|\omega| > 2{,}000\pi$. Amplitude modulation is performed to produce the signal

$$g(t) = x(t)\sin(2{,}000\pi t).$$

A proposed demodulation technique is illustrated in Figure P8.3 where $g(t)$ is the input, $y(t)$ is the output, and the ideal lowpass filter has cutoff frequency $2{,}000\pi$ and passband gain of 2. Determine $y(t)$.

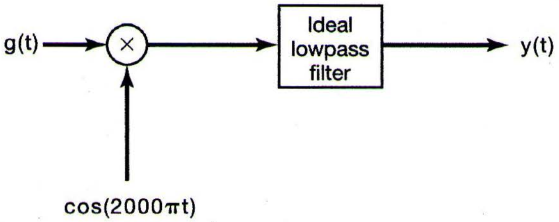


Figure P8.3

8.4. Suppose

$$x(t) = \sin 200\pi t + 2\sin 400\pi t$$

and

$$g(t) = x(t)\sin 400\pi t.$$

If the product $g(t)(\sin 400\pi t)$ is passed through an ideal lowpass filter with cutoff frequency 400π and passband gain of 2, determine the signal obtained at the output of the lowpass filter.

8.5. Suppose we wish to transmit the signal

$$x(t) = \frac{\sin 1{,}000\pi t}{\pi t}$$

using a modulator that creates the signal

$$w(t) = (x(t) + A)\cos(10{,}000\pi t).$$

Determine the largest permissible value of the modulation index m that would allow asynchronous demodulation to be used to recover $x(t)$ from $w(t)$. For this problem, you should assume that the maximum magnitude taken on by a side lobe of a sinc function occurs at the instant of time that is exactly halfway between the two zero-crossings enclosing the side lobe.

8.6. Assume that $x(t)$ is a signal whose Fourier transform $X(j\omega)$ is zero for $|\omega| > \omega_M$. The signal $g(t)$ may be expressed in terms of $x(t)$ as

$$g(t) = x(t)\cos\omega_c t - \left\{ x(t)\cos\omega_c t * \left(\frac{\sin\omega_c t}{\pi t}\right)\right\},$$

where $*$ denotes convolution and $\omega_c > \omega_M$. Determine the value of the constant A such that

$$x(t) = (g(t)\cos\omega_c t) * \frac{A\sin\omega_M t}{\pi t}.$$

8.7. An AM-SSB/SC system is applied to a signal $x(t)$ whose Fourier transform $X(j\omega)$ is zero for $|\omega| > \omega_M$. The carrier frequency ω_c used in the system is greater than ω_M. Let $g(t)$ denote the output of the system, assuming that only the *upper* sidebands are retained. Let $q(t)$ denote the output of the system, assuming that only the *lower* sidebands are retained. The system in Figure P8.7 is proposed for converting $g(t)$ into $q(t)$. How should the parameter ω_0 in the figure be related to ω_c? What should be the value of passband gain A?

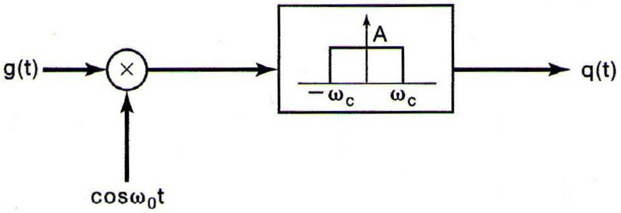


Figure P8.7

8.8. Consider the modulation system shown in Figure P8.8. The input signal $x(t)$ has a Fourier transform $X(j\omega)$ that is zero for $|\omega| > \omega_M$. Assuming that $\omega_c > \omega_M$, answer the following questions:

(a) Is $y(t)$ guaranteed to be real if $x(t)$ is real?
(b) Can $x(t)$ be recovered from $y(t)$?

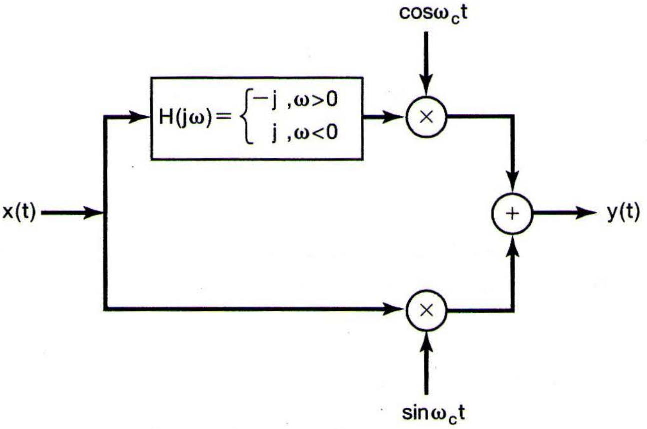


Figure P8.8

8.9. Two signals $x_1(t)$ and $x_2(t)$, each with a Fourier transform that is zero for $|\omega| > \omega_c$, are to be combined using frequency-division multiplexing. The AM-SSB/SC technique of Figure 8.21 is applied to each signal in a manner that retains the lower sidebands. The carrier frequencies used for $x_1(t)$ and $x_2(t)$ are ω_c and $2\omega_c$, respectively. The two modulated signals are then summed together to obtain the FDM signal $y(t)$.

(a) For what values of ω is $Y(j\omega)$ guaranteed to be zero?
(b) Specify the values of A and ω_0 so that

$$x_1(t) = \left[\left\{y(t) * \frac{\sin \omega_0 t}{\pi t}\right\} \cos \omega_0 t\right] * \frac{A \sin \omega_c t}{\pi t},$$

where $*$ denotes convolution.

8.10. A signal $x(t)$ is multiplied by the rectangular pulse train $c(t)$ shown in Figure P8.10.
(a) What constraint should be placed on $X(j\omega)$ to ensure that $x(t)$ can be recovered from the product $x(t)c(t)$ by using an ideal lowpass filter?
(b) Specify the cutoff frequency ω_c and the passband gain A of the ideal lowpass filter needed to recover $x(t)$ from $x(t)c(t)$. [Assume that $X(j\omega)$ satisfies the constraint determined in part (a).]

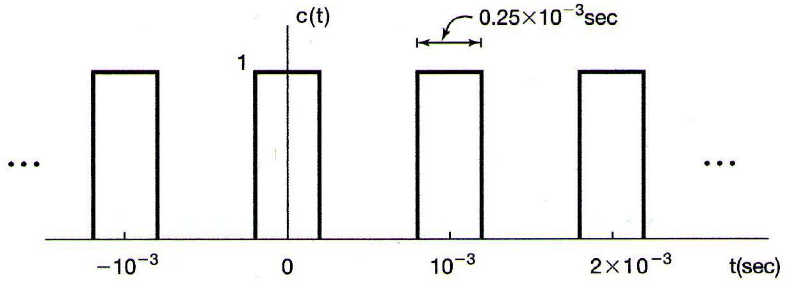


Figure P8.10

8.11. Let

$$c(t) = \sum_{k=-\infty}^{\infty} a_k e^{jk\omega_c t},$$

where $a_0 = 0$ and $a_1 \neq 0$, be a real-valued periodic signal. Also, let $x(t)$ be a signal with $X(j\omega) = 0$ for $|\omega| \geq \omega_c/2$. The signal $x(t)$ is used to modulate the carrier $c(t)$ to obtain

$$y(t) = x(t)c(t).$$

(a) Specify the passband and the passband gain of an ideal bandpass filter so that, with input $y(t)$, the output of the filter is

$$g(t) = (a_1 e^{j\omega_c t} + a_1^* e^{-j\omega_c t})x(t).$$

(b) If $a_1 = |a_1|e^{j\sphericalangle a_1}$, show that

$$g(t) = A\cos(\omega_c t + \phi)x(t),$$

and express A and ϕ in terms of $|a_1|$ and $\sphericalangle a_1$.

8.12. Consider a set of 10 signals $x_i(t)$, $i = 1, 2, 3, \ldots, 10$. Assume that each $x_i(t)$ has Fourier transform such that $X_i(j\omega) = 0$ for $|\omega| \geq 2{,}000\pi$. All 10 signals are to be time-division multiplexed after each is multiplied by a carrier $c(t)$ shown in Figure P8.12. If the period T of $c(t)$ is chosen to have the maximum allowable value, what is the largest value of Δ such that all 10 signals can be time-division multiplexed?

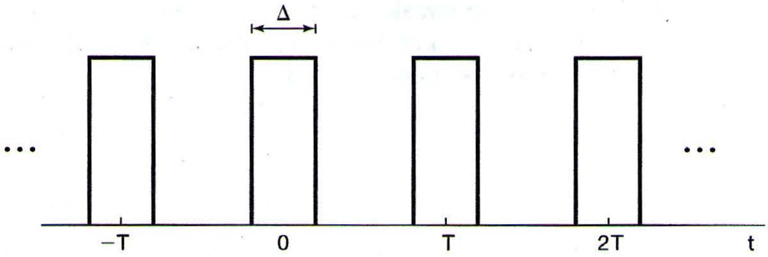


Figure P8.12

8.13. A class of popularly used pulses in PAM are those which have a *raised cosine* frequency response. The frequency response of one of the members of this class is

$$P(j\omega) = \begin{cases} \frac{1}{2}\left(1 + \cos\frac{\omega T_1}{2}\right), & 0 \leq |\omega| \leq \frac{2\pi}{T_1}, \\ 0, & \text{elsewhere} \end{cases}$$

where T_1 is the intersymbol spacing.

(a) Determine $p(0)$.

(b) Determine $p(kT_1)$, where $k = \pm 1, \pm 2, \ldots$.

8.14. Consider the frequency-modulated signal

$$y(t) = \cos(\omega_c t + m\cos\omega_m t),$$

where $\omega_c >> \omega_m$ and $m << \pi/2$. Specify an approximation to $Y(j\omega)$ for $\omega > 0$.

8.15. For what values of ω_0 in the range $-\pi < \omega_0 \leq \pi$ is amplitude modulation with carrier $e^{j\omega_0 n}$ equivalent to amplitude modulation with carrier $\cos\omega_0 n$?

8.16. Suppose $x[n]$ is a real-valued discrete-time signal whose Fourier transform $X(e^{j\omega})$ has the property that

$$X(e^{j\omega}) = 0 \quad \text{for } \frac{\pi}{8} \leq \omega \leq \pi.$$

We use $x[n]$ to modulate a sinusoidal carrier $c[n] = \sin(5\pi/2)n$ to produce

$$y[n] = x[n]c[n].$$

Determine the values of ω in the range $0 \leq \omega \leq \pi$ for which $Y(e^{j\omega})$ is guaranteed to be zero.

8.17. Consider an arbitrary finite-duration signal $x[n]$ with Fourier transform $X(e^{j\omega})$. We generate a signal $g[n]$ through insertion of zero-valued samples:

$$g[n] = x_{(4)}[n] = \begin{cases} x[n/4], & n = 0, \pm 4, \pm 8, \pm 12, \ldots \\ 0, & \text{otherwise} \end{cases}.$$

The signal $g[n]$ is passed through an ideal lowpass filter with cutoff frequency $\pi/4$ and passband gain of unity to produce a signal $q[n]$. Finally, we obtain

$$y[n] = q[n]\cos\left(\frac{3\pi}{4}n\right).$$

For what values of ω is $Y(e^{j\omega})$ guaranteed to be zero?

8.18. Let $x[n]$ be a real-valued discrete-time signal whose Fourier transform $X(e^{j\omega})$ is zero for $\omega \geq \pi/4$. We wish to obtain a signal $y[n]$ whose Fourier transform has the property that, in the interval $-\pi < \omega \leq \pi$,

$$Y(e^{j\omega}) = \begin{cases} X(e^{j(\omega-\frac{\pi}{2})}), & \frac{\pi}{2} < \omega \leq \frac{3\pi}{4} \\ X(e^{j(\omega+\frac{\pi}{2})}), & -\frac{3\pi}{4} < \omega \leq -\frac{\pi}{2} \\ 0, & \text{otherwise} \end{cases}.$$

The system in Figure P8.18 is proposed for obtaining $y[n]$ from $x[n]$. Determine constraints that the frequency response $H(e^{j\omega})$ of the filter in the figure must satisfy for the proposed system to work.

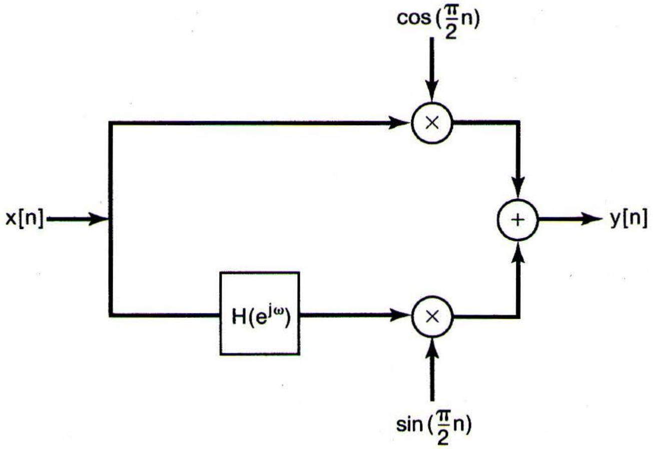


Figure P8.18

8.19. Consider 10 arbitrary real-valued signals $x_i[n]$, $i = 1, 2, \ldots, 10$. Suppose each $x_i[n]$ is upsampled by a factor of N, and then sinusoidal amplitude modulation is applied to it with carrier frequency $\omega_i = i\pi/10$. Determine the value of N which would guarantee that all 10 modulated signals can be summed together to yield an FDM signal $y[n]$ from which each $x_i[n]$ can be recovered.

8.20. Let $v_1[n]$ and $v_2[n]$ be two discrete-time signals obtained through the sampling (without aliasing) of continuous-time signals. Let

$$y[n] = \hat{v}_1[n] + \hat{v}_2[n-1]$$

be a TDM signal, where, for $i = 1, 2$,

$$\hat{v}_i[n] = \begin{cases} v_i\left[\frac{n}{2}\right], & n = 0, \pm 2, \pm 4, \pm 6, \ldots \\ 0, & \text{otherwise} \end{cases}.$$

The signal $y[n]$ is processed by the system S depicted in Figure P8.20 to obtain a signal $g[n]$. For the two filters used in S,

$$H_0(e^{j\omega}) = \begin{cases} 1, & |\omega| \leq \frac{\pi}{2} \\ 0, & \frac{\pi}{2} < \omega \leq \pi \end{cases}.$$

Determine the signal $p[n]$ used in S such that $g[n]$ represents frequency-division multiplexing of $v_1[n]$ and $v_2[n]$.

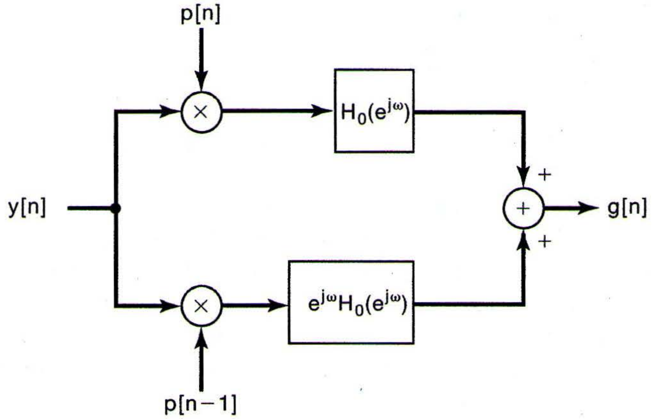


Figure P8.20

BASIC PROBLEMS

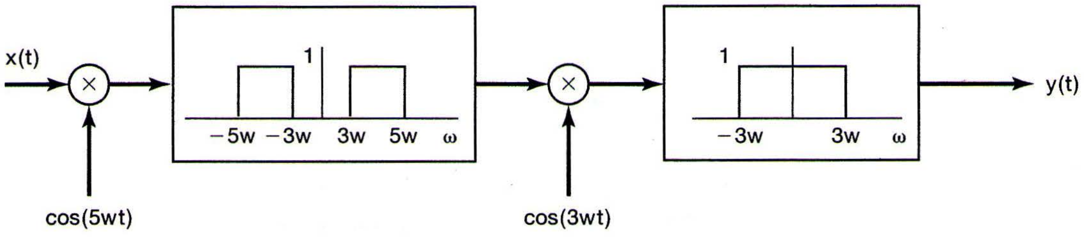
8.21. In Sections 8.1 and 8.2, we analyzed the sinusoidal amplitude modulation and demodulation system of Figure 8.8, assuming that the phase θ_c of the carrier signal was zero.

(a) For the more general case of arbitrary phase θ_c in the figure, show that the signal in the demodulation system can be expressed as

$$w(t) = \frac{1}{2}x(t) + \frac{1}{2}x(t)\cos(2\omega_c t + 2\theta_c).$$

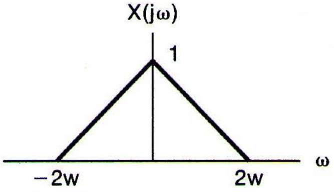
(b) If $x(t)$ has a spectrum that is zero for $|\omega| \geq \omega_M$, determine the relationships required among w_{co} [the cutoff frequency of the ideal lowpass filter in Figure 8.8(b)], ω_c (the carrier frequency), and ω_M so that the output of the lowpass filter is proportional to $x(t)$. Does your answer depend on the carrier phase θ_c?

8.22. In Figure P8.22(a), a system is shown with input signal $x(t)$ and output signal $y(t)$. The input signal has the Fourier transform $X(j\omega)$ shown in Figure P8.22(b). Determine and sketch $Y(j\omega)$, the spectrum of $y(t)$.



(a)

Figure P8.22



(b)

Figure P8.22 *Continued*

8.23. In Section 8.2, we discussed the effect of a loss of synchronization in phase between the carrier signals in the modulator and demodulator in sinusoidal amplitude modulation. We showed that the output of the demodulation is attenuated by the cosine of the phase difference, and in particular, when the modulator and demodulator have a phase difference of $\pi/2$, the demodulator output is zero. As we demonstrate in this problem, it is also important to have *frequency* synchronization between the modulator and demodulator.

Consider the amplitude modulation and demodulation systems in Figure 8.8 with $\theta_c = 0$ and with a change in the *frequency* of the demodulator carrier so that

$$w(t) = y(t)\cos\omega_d t,$$

where

$$y(t) = x(t)\cos\omega_c t.$$

Let us denote the difference in frequency between the modulator and demodulator as $\Delta\omega$ (i.e., $\omega_d - \omega_c = \Delta\omega$). Also, assume that $x(t)$ is band limited with $X(j\omega) = 0$ for $|\omega| \geq \omega_M$, and assume that the cutoff frequency w_{co} of the lowpass filter in the demodulator satisfies the inequality

$$\omega_M + \Delta\omega < \mathrm{w}_{co} < 2\omega_c + \Delta\omega - \omega_M.$$

(a) Show that the output of the lowpass filter in the demodulator is proportional to $x(t)\cos(\Delta\omega t)$.

(b) If the spectrum of $x(t)$ is that shown in Figure P8.23, sketch the spectrum of the output of the demodulator.

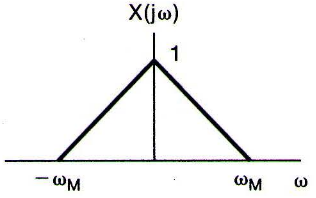


Figure P8.23

8.24. Figure P8.24 shows a system to be used for sinusoidal amplitude modulation, where $x(t)$ is band limited with maximum frequency ω_M, so that $X(j\omega) = 0$, $|\omega| > \omega_M$.

As indicated, the signal $s(t)$ is a periodic impulse train with period T and with an offset from $t = 0$ of Δ. The system $H(j\omega)$ is a bandpass filter.

(a) With $\Delta = 0$, $\omega_M = \pi/2T$, $\omega_l = \pi/T$, and $\omega_h = 3\pi/T$, show that $y(t)$ is proportional to $x(t)\cos\omega_c t$, where $\omega_c = 2\pi/T$.

(b) If ω_M, ω_l, and ω_h are the same as given in part (a), but Δ is not necessarily zero, show that $y(t)$ is proportional to $x(t)\cos(\omega_c t + \theta_c)$, and determine ω_c and θ_c as a function of T and Δ.

(c) Determine the maximum allowable value of ω_M relative to T such that $y(t)$ is proportional to $x(t)\cos(\omega_c t + \theta_c)$.

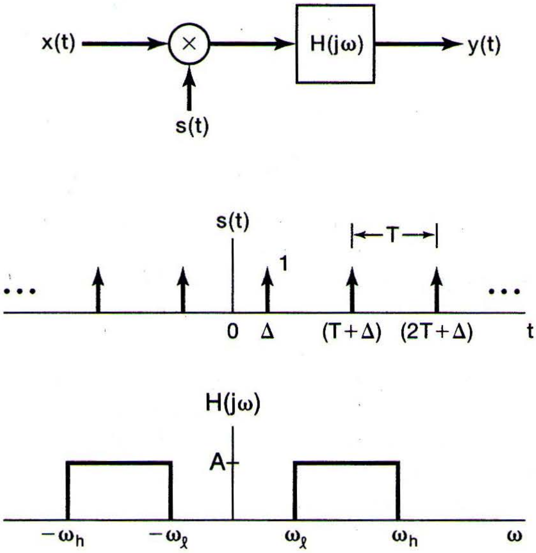


Figure P8.24

8.25. A commonly used system to maintain privacy in voice communication is a *speech scrambler.* As illustrated in Figure P8.25(a), the input to the system is a normal speech signal $x(t)$ and the output is the scrambled version $y(t)$. The signal $y(t)$ is transmitted and then unscrambled at the receiver.

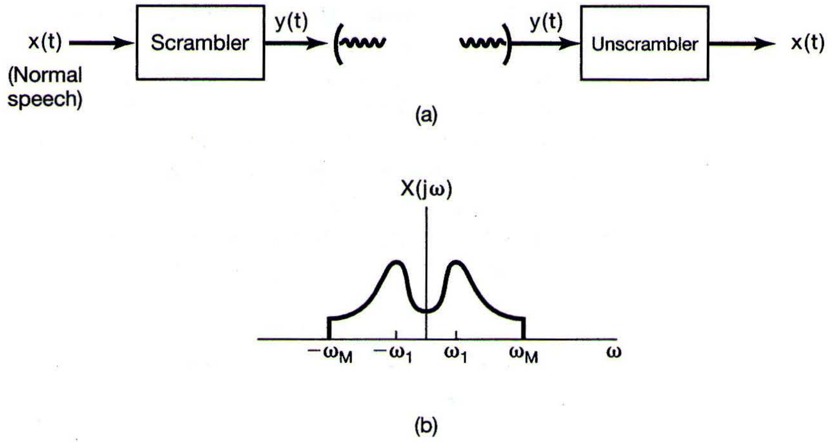
We assume that all inputs to the scrambler are real and band limited to the frequency ω_M; that is, $X(j\omega) = 0$ for $|\omega| > \omega_M$. Given any such input, our proposed scrambler permutes different bands of the spectrum of the input signal. In addition, the output signal is real and band limited to the same frequency band; that is, $Y(j\omega) = 0$ for $|\omega| > \omega_M$. The specific algorithm for the scrambler is

$$Y(j\omega) = X(j(\omega - \omega_M)), \quad \omega > 0,$$
$$Y(j\omega) = X(j(\omega + \omega_M)), \quad \omega < 0.$$

(a) If $X(j\omega)$ is given by the spectrum shown in Figure P8.25(b), sketch the spectrum of the scrambled signal $y(t)$.

(b) Using amplifiers, multipliers, adders, oscillators, and whatever ideal filters you find necessary, draw the block diagram for such an ideal scrambler.

(c) Again using amplifiers, multipliers, adders, oscillators, and ideal filters, draw a block diagram for the associated unscrambler.


x(t)
(Normal speech)
Scrambler
y(t)
y(t)
Unscrambler
x(t)
X(jω)
−ωM
−ω1
ω1
ωM
ω

(a)

(b)

Figure P8.25

8.26. In Section 8.2.2, we discussed the use of an envelope detector for asynchronous demodulation of an AM signal of the form $y(t) = [x(t) + A]\cos(\omega_c t + \theta_c)$. An alternative demodulation system, which also does not require phase synchronization, but does require frequency synchronization, is shown in block diagram form in Figure P8.26. The lowpass filters both have a cutoff frequency of ω_c. The signal $y(t) = [x(t) + A]\cos(\omega_c t + \theta_c)$, with θ_c constant but unknown. The signal $x(t)$ is band limited with $X(j\omega) = 0,\ |\omega| > \omega_M$, and with $\omega_M < \omega_c$. As we required for the use of the envelope detector, $x(t) + A > 0$ for all t.

Show that the system in Figure P8.26 can be used to recover $x(t)$ from $y(t)$ without knowledge of the modulator phase θ_c.

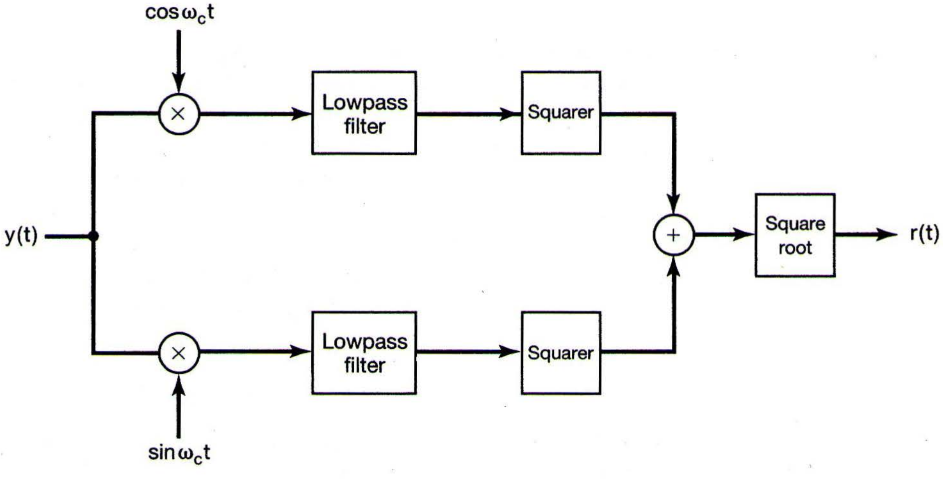

cos ωct
×
Lowpass filter
Squarer
y(t)
+
Square root
r(t)
×
Lowpass filter
Squarer
sin ωct

Figure P8.26

8.27. As discussed in Section 8.2.2, asynchronous modulation-demodulation requires the injection of the carrier signal so that the modulated signal is of the form

$$y(t) = [A + x(t)]\cos(\omega_c t + \theta_c), \tag{P8.27–1}$$

where $A + x(t) > 0$ for all t. The presence of the carrier means that more transmitter power is required, representing an inefficiency.

(a) Let $x(t) = \cos\omega_M t$ with $\omega_M < \omega_c$ and $A + x(t) > 0$. For a periodic signal $y(t)$ with period T, the average power over time is defined as $P_y = (1/T)\int_T y^2(t)\,dt$. Determine and sketch P_y for $y(t)$ in eq. (P8.27–1). Express your answer as a function of the modulation index m, defined as the maximum absolute value of $x(t)$ divided by A.

(b) The efficiency of transmission of an amplitude-modulated signal is defined to be the ratio of the power in the sidebands of the signal to the total power in the signal. With $x(t) = \cos\omega_M t$, and with $\omega_M < \omega_c$ and $A + x(t) > 0$, determine and sketch the efficiency of the modulated signal as a function of the modulation index m.

8.28. In Section 8.4 we discussed the implementation of single-sideband modulation using 90° phase-shift networks, and in Figures 8.21 and 8.22 we specifically illustrated the system and associated spectra required to retain the lower sidebands.

Figure P8.28(a) shows the corresponding system required to retain the upper sidebands.

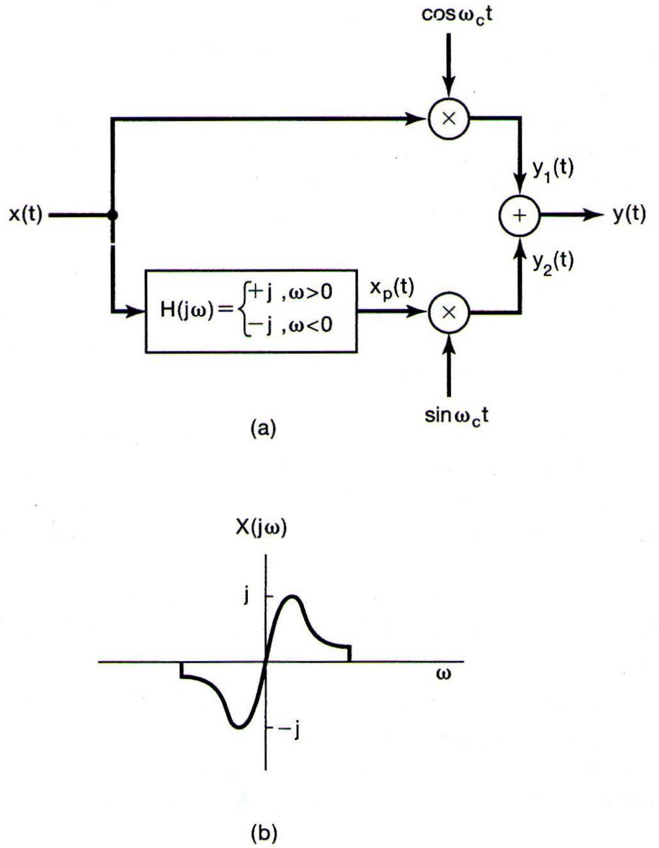


Figure P8.28

(a) With the same $X(j\omega)$ illustrated in Figure 8.22, sketch $Y_1(j\omega)$, $Y_2(j\omega)$, and $Y(j\omega)$ for the system in Figure P8.28(a), and demonstrate that only the upper sidebands are retained.

(b) For $X(j\omega)$ imaginary, as illustrated in Figure P8.28(b), sketch $Y_1(j\omega)$, $Y_2(j\omega)$, and $Y(j\omega)$ for the system in Figure P8.28(a), and demonstrate that, for this case also, only the upper sidebands are retained.

8.29. Single-sideband modulation is commonly used in point-to-point voice communication. It offers many advantages, including effective use of available power, conservation of bandwidth, and insensitivity to some forms of random fading in the channel. In double-sideband suppressed carrier (DSB/SC) systems the spectrum of the modulating signal appears in its entirety in two places in the transmitted spectrum. Single-sideband modulation eliminates this redundancy, thus conserving bandwidth and increasing the signal-to-noise ratio within the remaining portion of the spectrum that is transmitted.

In Figure P8.29(a), two systems for generating an amplitude-modulated single-sideband signal are shown. The system on the top can be used to generate a single-sideband signal for which the lower sideband is retained, and the system on the bottom can produce a single-sideband signal for which the upper sideband is retained.

(a) For $X(j\omega)$ as shown in Figure P8.29(b), determine and sketch $S(j\omega)$, the Fourier transform of the lower sideband modulated signal, and $R(j\omega)$, the Fourier transform of the upper sideband modulated signal. Assume that $\omega_c > \omega_3$.

The upper sideband modulation scheme is particularly useful with voice communication, as any real filter has a finite transition region for the cutoff (i.e., near ω_c). This region can be accommodated with negligible distortion, since the voice signal does not have any significant energy near $\omega = 0$ (i.e., for $|\omega| < \omega_1 = 2\pi \times 40$ Hz).

(b) Another procedure for generating a single-sideband signal is termed the *phase-shift method* and is illustrated in Figure P8.29(c). Show that the single-sideband signal generated is proportional to that generated by the lower sideband modulation scheme of Figure P8.29(a) [i.e., $p(t)$ is proportional to $s(t)$].

(c) All three AM-SSB signals can be demodulated using the scheme shown on the right-hand side of Figure P8.29(a). Show that, whether the received signal is $s(t)$, $r(t)$, or $p(t)$, as long as the oscillator at the receiver is in phase with oscillators at the transmitter, and $w = \omega_c$, the output of the demodulator is $x(t)$.

The distortion that results when the oscillator is not in phase with the transmitter, called *quadrature distortion,* can be particularly troublesome in data communication.

8.30. Amplitude modulation with a pulse-train carrier may be modeled as in Figure P8.30(a). The output of the system is $q(t)$.

(a) Let $x(t)$ be a band-limited signal [i.e., $X(j\omega) = 0$, $|\omega| \geq \pi/T$], as shown in Figure P8.30(b). Determine and sketch $R(j\omega)$ and $Q(j\omega)$.

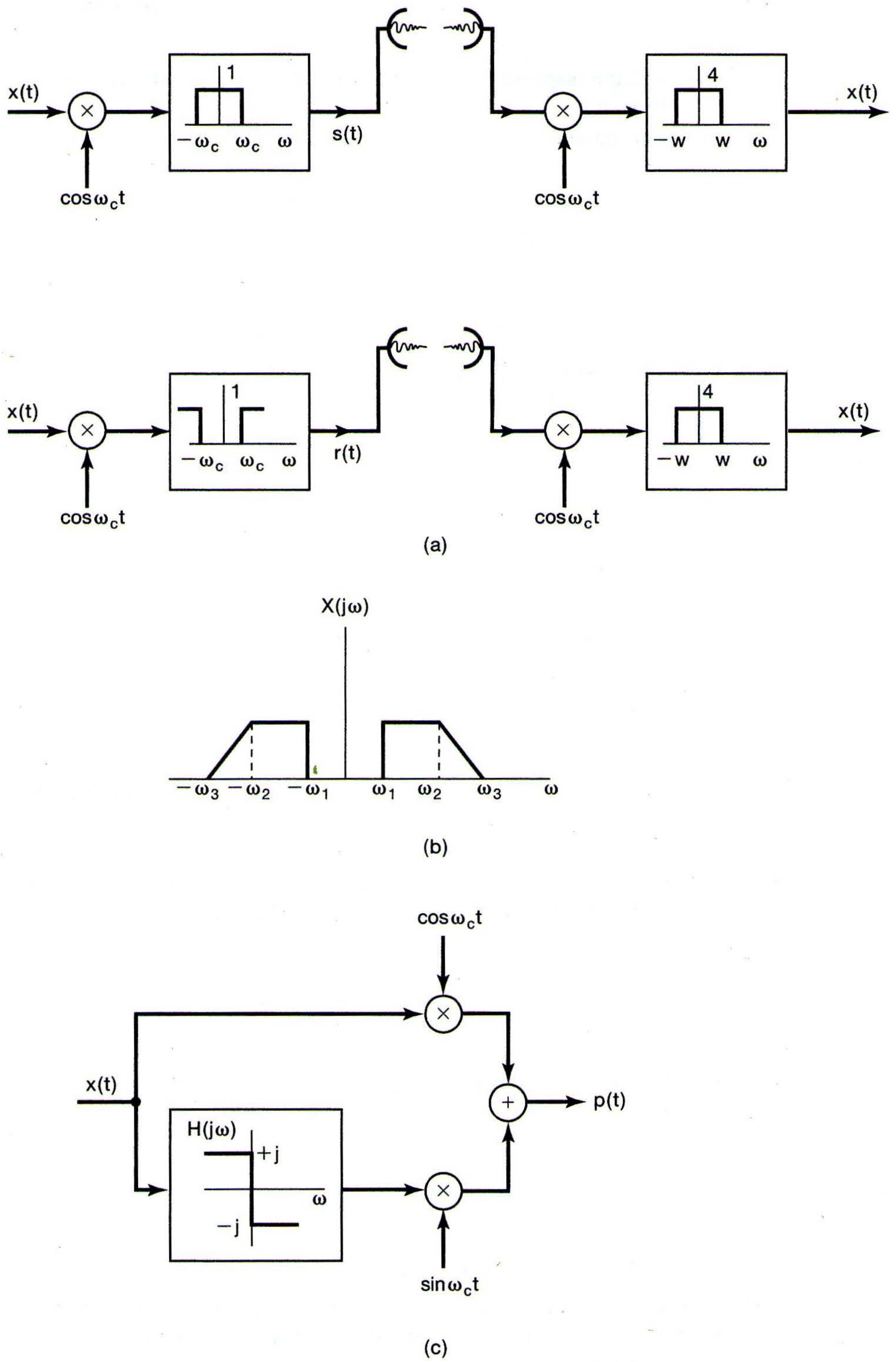
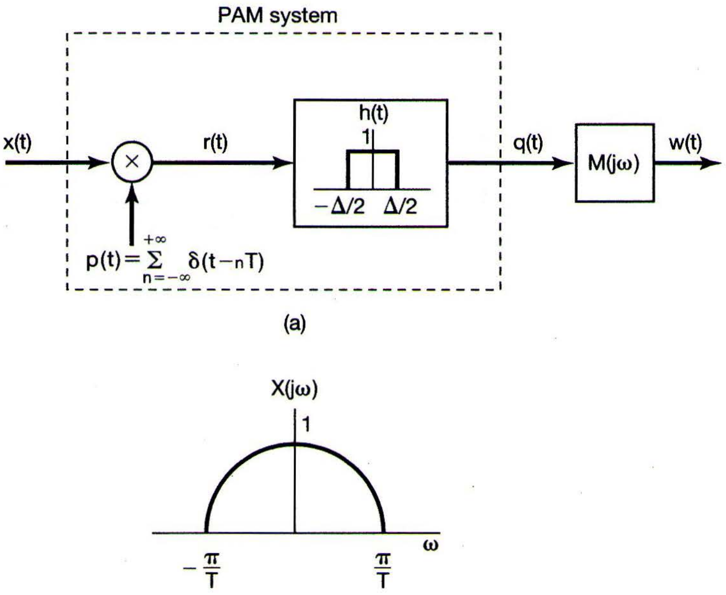

x(t)
cos ω_c t
1
−ω_c
ω_c
ω
s(t)
4
−W
W
ω
cos ω_c t
x(t)
x(t)
cos ω_c t
1
−ω_c
ω_c
ω
r(t)
4
−W
W
ω
cos ω_c t
x(t)
(a)
X(jω)
−ω_3
−ω_2
−ω_1
ω_1
ω_2
ω_3
ω
(b)
cos ω_c t
x(t)
H(jω)
+j
−j
ω
p(t)
sin ω_c t
(c)

Figure P8.29

(b) Find the maximum value of Δ such that $w(t) = x(t)$ with an appropriate filter $M(j\omega)$.

(c) Determine and sketch the compensating filter $M(j\omega)$ such that $w(t) = x(t)$.


PAM system
x(t)
r(t)
h(t)
1
−Δ/2
Δ/2
q(t)
M(jω)
w(t)
$p(t)=\sum_{n=-\infty}^{+\infty}\delta(t-nT)$
X(jω)
1
ω
$-\frac{\pi}{T}$
$\frac{\pi}{T}$

(a)

(b)

Figure P8.30

8.31. Let $x[n]$ be a discrete-time signal with spectrum $X(e^{j\omega})$, and let $p(t)$ be a continuous-time pulse function with spectrum $P(j\omega)$. We form the signal

$$y(t) = \sum_{n=-\infty}^{+\infty} x[n]p(t-n).$$

(a) Determine the spectrum $Y(j\omega)$ in terms of $X(e^{j\omega})$ and $P(j\omega)$.

(b) If

$$p(t) = \begin{cases} \cos 8\pi t, & 0 \le t \le 1 \\ 0, & \text{elsewhere} \end{cases},$$

determine $P(j\omega)$ and $Y(j\omega)$.

8.32. Consider a discrete-time signal $x[n]$ with Fourier transform shown in Figure P8.32(a). The signal is amplitude modulated by a sinusoidal sequence, as indicated in Figure P8.32(b).

(a) Determine and sketch $Y(e^{j\omega})$, the Fourier transform of $y[n]$.

(b) A proposed demodulation system is shown in Figure P8.32(c). For what value of θ_c, ω_{lp}, and G will $\hat{x}[n] = x[n]$? Are any restrictions on ω_c and ω_{lp} necessary to guarantee that $x[n]$ is recoverable from $y[n]$?

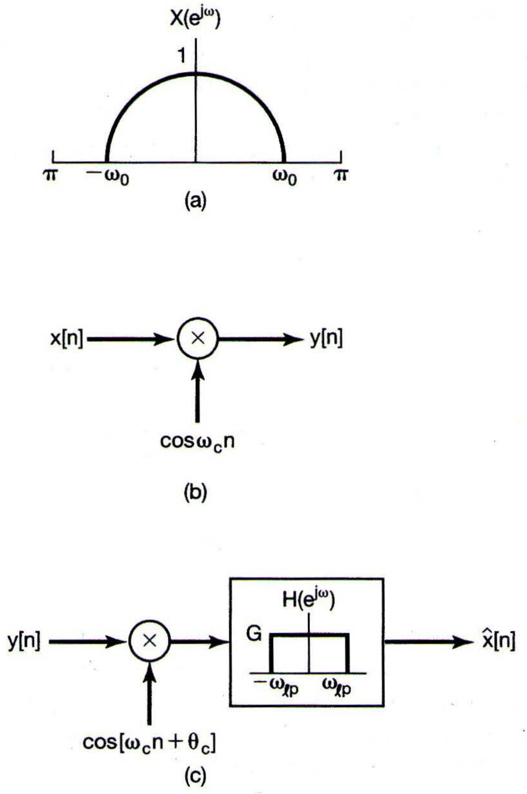


Figure P8.32

8.33. Let us consider the frequency-division multiplexing of discrete-time signals $x_i[n]$, $i = 0, 1, 2, 3$. Furthermore, each $x_i[n]$ potentially occupies the entire frequency band $(-\pi < \omega < \pi)$. The sinusoidal modulation of upsampled versions of each of these signals may be carried out by using either double-sideband techniques or single-sideband techniques.

(a) Suppose each signal $x_i[n]$ is appropriately upsampled and then modulated with $\cos[i(\pi/4)n]$. What is the minimum amount of upsampling that must be carried out on each $x_i[n]$ in order to ensure that the spectrum of the FDM signal does not have any aliasing?

(b) If the upsampling of each $x_i[n]$ is restricted to be by a factor of 4, how would you use single-sideband techniques to ensure that the FDM signal does not have any aliasing? *Hint:* See Problem 8.17.

ADVANCED PROBLEMS

8.34. In discussing amplitude modulation systems, modulation and demodulation were carried out through the use of a multiplier. Since multipliers are often difficult to implement, many practical systems use a nonlinear element. In this problem, we illustrate the basic concept.

In Figure P8.34, we show one such nonlinear system for amplitude modulation. The system consists of squaring the *sum* of the modulating signal and the carrier and then bandpass filtering to obtain the amplitude-modulated signal.

Assume that $x(t)$ is band limited, so that $X(j\omega) = 0, \ |\omega| > \omega_M$. Determine the bandpass filter parameters A, ω_l, and ω_h such that $y(t)$ is an amplitude-modulated version of $x(t)$ [i.e., such that $y(t) = x(t)\cos\omega_c t$]. Specify the necessary constraints, if any, on ω_c and ω_M.

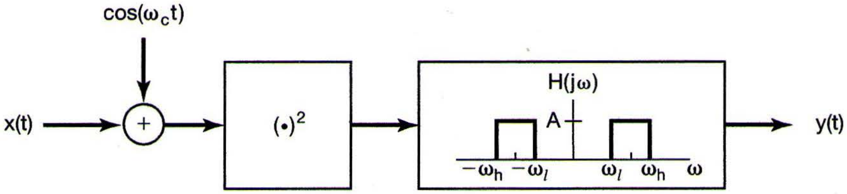
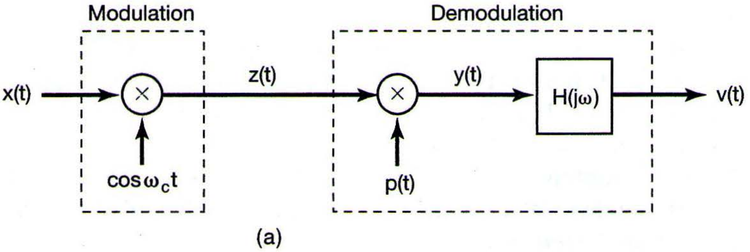


Figure P8.34

8.35. The modulation-demodulation scheme proposed in this problem is similar to sinusoidal amplitude modulation, except that the demodulation is done with a square wave with the same zero-crossings as $\cos\omega_c t$. The system is shown in Figure P8.35(a); the relation between $\cos\omega_c t$ and $p(t)$ is shown in Figure P8.35(b). Let the input signal $x(t)$ be a band-limited signal with maximum frequency $\omega_M < \omega_c$, as shown in Figure P8.35(c).

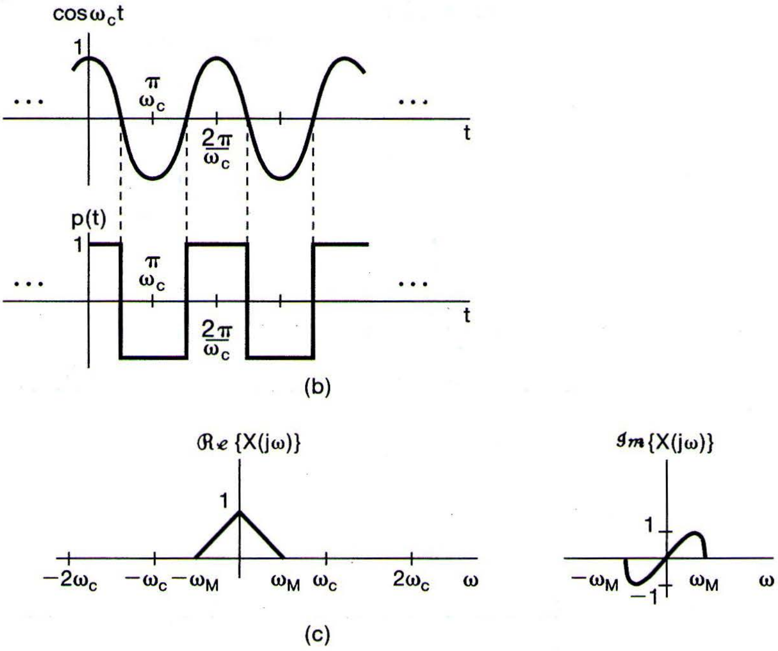
(a) Sketch and dimension the real and imaginary parts of $Z(j\omega)$, $P(j\omega)$, and $Y(j\omega)$, the Fourier transforms of $z(t)$, $p(t)$, and $y(t)$, respectively.

(b) Sketch and dimension a filter $H(j\omega)$ so that $v(t) = x(t)$.



(a)

Figure P8.35



(b)

(c)

Figure P8.35 *Continued*

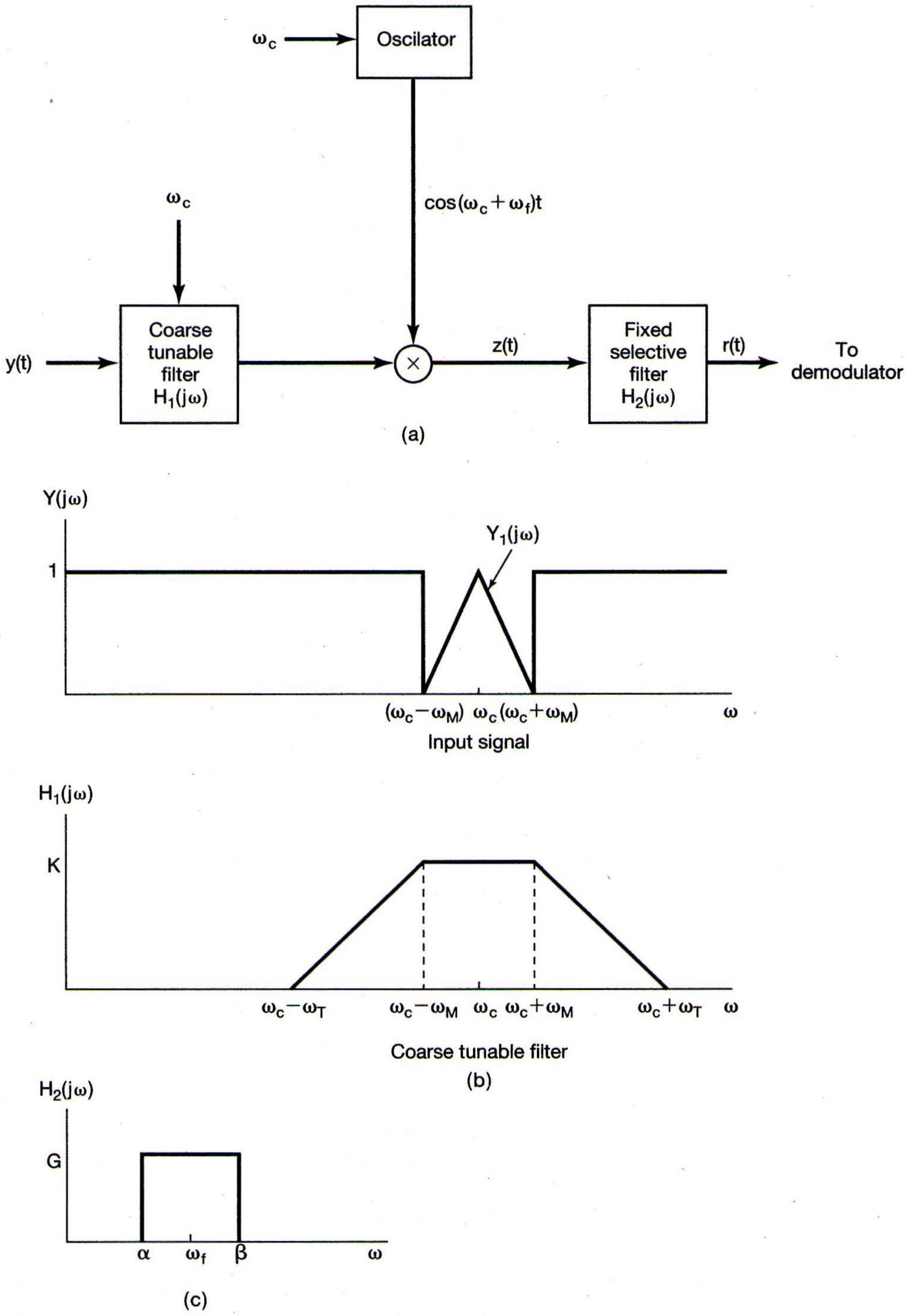
8.36. The accurate demultiplexing-demodulation of radio and television signals is generally performed using a system called the superheterodyne receiver, which is equivalent to a tunable filter. The basic system is shown in Figure P8.36(a).

(a) The input signal $y(t)$ consists of the superposition of many amplitude-modulated signals that have been multiplexed using frequency-division multiplexing, so that each signal occupies a different frequency channel. Let us consider one such channel that contains the amplitude-modulated signal $y_1(t) = x_1(t)\cos\omega_c t$, with spectrum $Y_1(j\omega)$ as depicted at the top of Figure P8.36(b). We want to demultiplex and demodulate $y_1(t)$ to recover the modulating signal $x_1(t)$, using the system of Figure P8.36(a). The coarse tunable filter has the spectrum $H_1(j\omega)$ shown at the bottom of Figure P8.36(b). Determine the spectrum $Z(j\omega)$ of the input signal to the fixed selective filter $H_2(j\omega)$. Sketch and label $Z(j\omega)$ for $\omega > 0$.

(b) The fixed frequency-selective filter is a bandpass type centered around the fixed frequency ω_f, as shown in Figure P8.36(c). We would like the output of the filter with spectrum $H_2(j\omega)$ to be $r(t) = x_1(t)\cos\omega_f t$. In terms of ω_c and ω_M, what constraint must ω_T satisfy to guarantee that an undistorted spectrum of $x_1(t)$ is centered around $\omega = \omega_f$?

(c) What must G, α, and β be in Figure P8.36(c) so that $r(t) = x_1(t)\cos\omega_f t$?

8.37. The following scheme has been proposed to perform amplitude modulation: The input signal $x(t)$ is added to the carrier signal $\cos\omega_c t$ and then put through a nonlinear



ω_c
Oscilator
$\cos(\omega_c+\omega_f)t$
ω_c
y(t)
Coarse tunable filter $H_1(j\omega)$
z(t)
Fixed selective filter $H_2(j\omega)$
r(t)
To demodulator
(a)
$Y(j\omega)$
$Y_1(j\omega)$
1
$(\omega_c-\omega_M)$ ω_c $(\omega_c+\omega_M)$
ω
Input signal
$H_1(j\omega)$
K
$\omega_c-\omega_T$
$\omega_c-\omega_M$
ω_c
$\omega_c+\omega_M$
$\omega_c+\omega_T$
ω
Coarse tunable filter
(b)
$H_2(j\omega)$
G
α
ω_f
β
ω
(c)

Figure P8.36

device, so that the output $z(t)$ is related to the input by

$$z(t) = e^{y(t)} - 1,$$
$$y(t) = x(t) + \cos \omega_c t.$$

This is illustrated in Figure P8.37(a). Such a nonlinear relation can be implemented through the current-voltage characteristics of a diode, where, with $i(t)$ and $v(t)$ the diode and current and voltage, respectively,

$$i(t) = I_0 e^{av(t)} - 1 \quad (a \text{ real}).$$

To study the effects of nonlinearity, we can examine the spectrum of $z(t)$ and how it relates to $X(j\omega)$ and ω_c. To accomplish this, we use the power series for e^y, which is

$$e^y = 1 + y + \frac{1}{2}y^2 + \frac{1}{6}y^3 + \cdots.$$

(a) If the spectrum of $x(t)$ is given by Figure P8.37(b), and if $\omega_c = 100\omega_1$, sketch and label $Z(j\omega)$, the spectrum of $z(t)$, using the first four terms in the power series for e^y.

(b) The bandpass filter has parameters as shown in Figure P8.37(c). Determine the range of α and the range of β such that $r(t)$ is an amplitude-modulated version of $x(t)$.

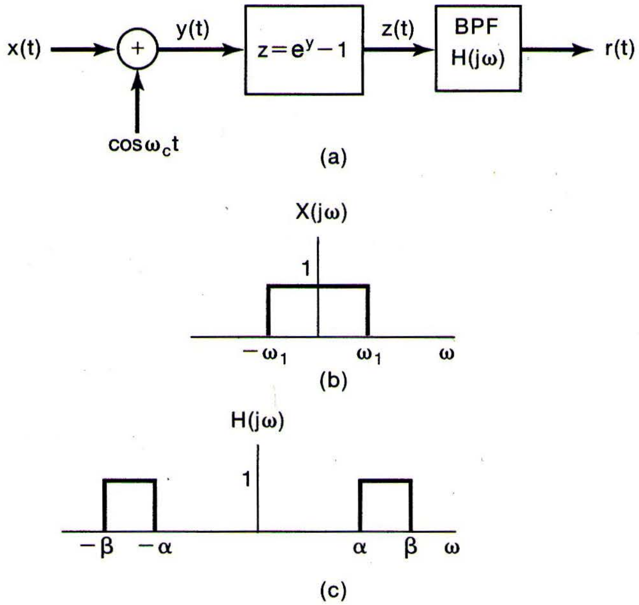


Figure P8.37

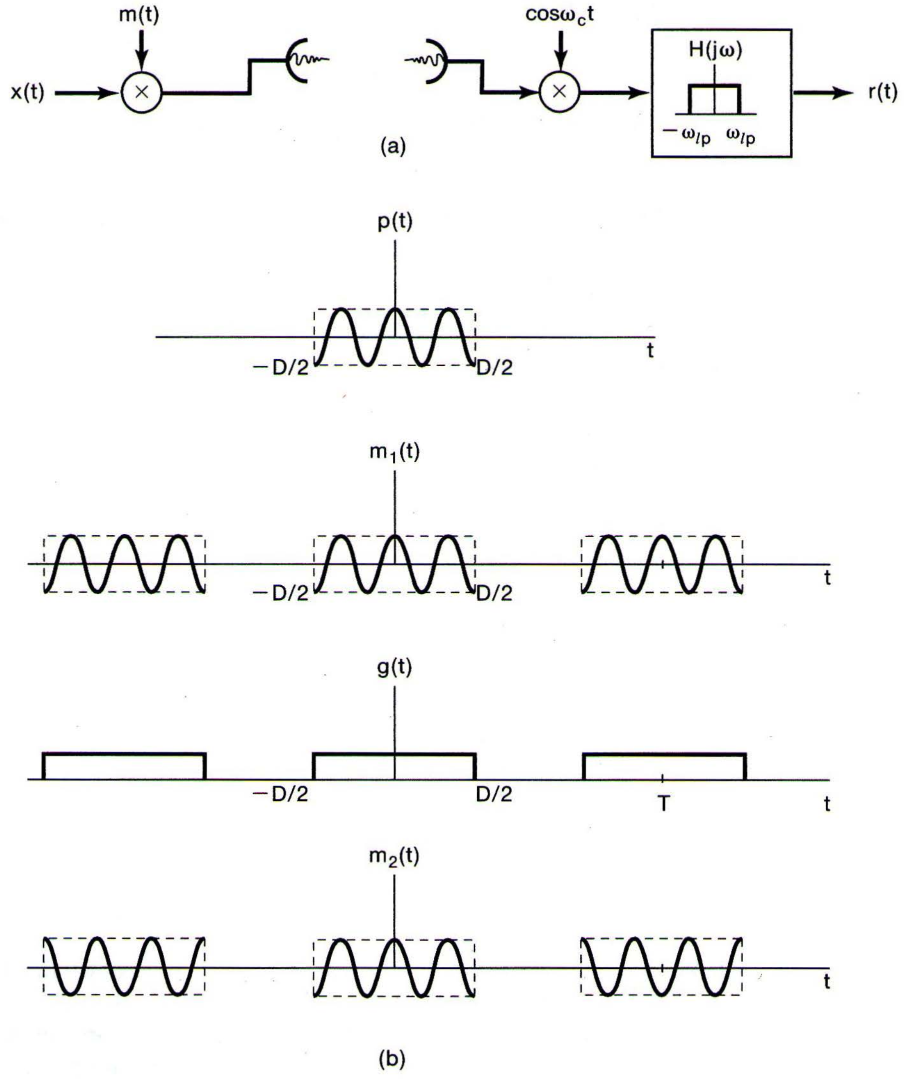
8.38. In Figure P8.38(a), a communication system is shown that transmits a band-limited signal $x(t)$ as periodic bursts of high-frequency energy. Assume that $X(j\omega) = 0$ for $|\omega| > \omega_M$. Two possible choices, $m_1(t)$ and $m_2(t)$, are considered for the modulating signal $m(t)$. $m_1(t)$ is a periodic train of sinusoidal pulses, each of duration D, as shown in Figure P8.38(b). That is,

$$m_1(t) = \sum_{k=-\infty}^{\infty} p(t - kT),$$

where

$$p(t) = \begin{cases} \cos\omega_c t, & |t| < (D/2) \\ 0, & |t| > (D/2) \end{cases}.$$

$m_2(t)$ is $\cos\omega_c t$ periodically blanked or gated; that is, $m_2(t) = g(t)\cos\omega_c t$, where $g(t)$ is as shown in Figure P8.38(b).


m(t)
x(t)
cosω_ct
H(jω)
−ω_lp
ω_lp
r(t)
p(t)
−D/2
D/2
t
m_1(t)
g(t)
T
m_2(t)

(a)

(b)

Figure P8.38

The following relationships between the parameters T, D, ω_c, and ω_M are assumed:

$$D < T,$$

$$\omega_c >> \frac{2\pi}{D},$$

$$\frac{2\pi}{T} > 2\omega_M.$$

Also, assume that $[\sin(x)]/x$ is negligible for $x >> 1$.

Determine whether, for some choice of ω_{lp}, either $m_1(t)$ or $m_2(t)$ will result in a demodulated signal $x(t)$. For each case in which your answer is yes, determine an acceptable range for ω_{lp}.

8.39. Suppose we wish to communicate one of two possible messages: message m_0 or message m_1. To do so, we will send a burst of one of two frequencies over a time interval of length T. Note that T is independent of which message is being transmitted. For message m_0 we will send $\cos \omega_0 t$, and for message m_1 we will send $\cos \omega_1 t$. Thus, a burst $b(t)$ will look as shown in Figure P8.39(a). Such a communi-

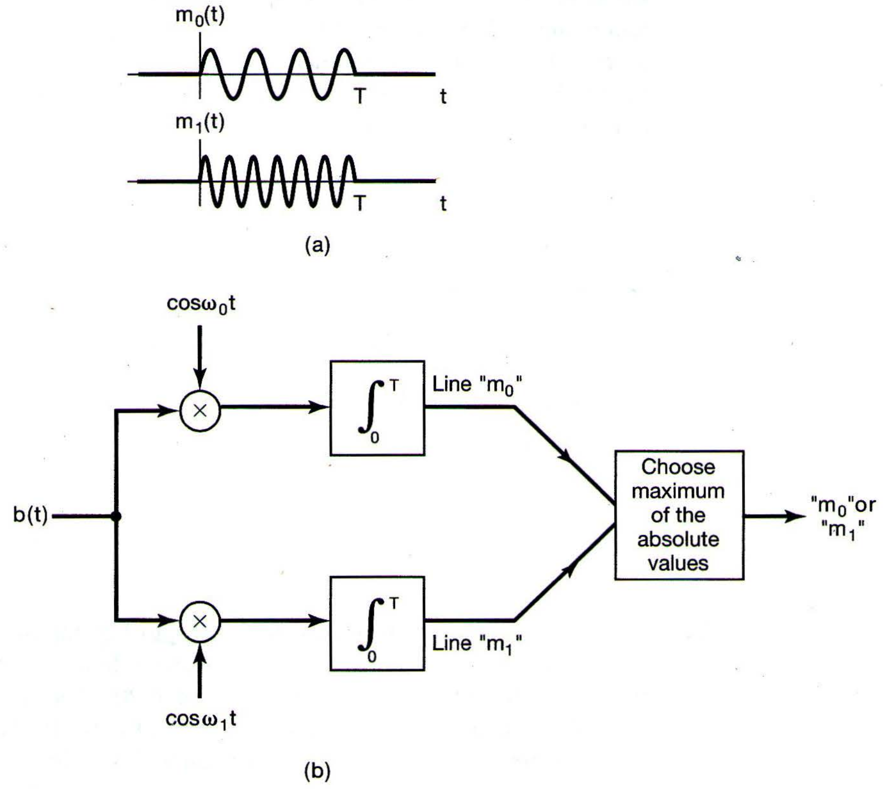

m0(t)
T
t
m1(t)
T
t
(a)
cosω0t
T
0
Line "m0"
Choose maximum of the absolute values
b(t)
"m0"or "m1"
T
0
Line "m1"
cosω1t
(b)

Figure P8.39

cation system is called *frequency shift keying* (FSK). When the burst of frequency $b(t)$ is received, we wish to determine whether it represents message m_0 or message m_1. To accomplish this, we do as illustrated in Figure P8.39(b).

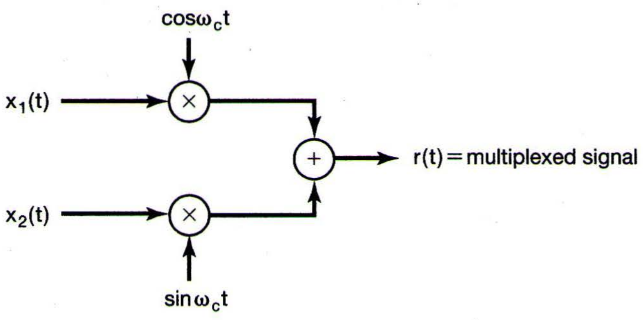
(a) Show that the maximum difference between the absolute values of the two lines in Figure P8.39(b) occurs when $\cos \omega_0 t$ and $\cos \omega_1 t$ have the relationship

$$\int_0^T \cos \omega_0 t \cos \omega_1 t \, dt = 0.$$

(b) Is it possible to choose ω_0 and ω_1 such that there is *no* interval of length T for which

$$\int_0^T \cos \omega_0 t \cos \omega_1 t \, dt = 0?$$

8.40. In Section 8.3, we discussed the use of sinusoidal modulation for frequency-division multiplexing whereby several signals are shifted into different frequency bands and then summed for simultaneous transmission. In the current problem, we explore another multiplexing concept referred to as *quadrature multiplexing*. In this multiplexing procedure, two signals can be transmitted simultaneously in the same frequency band if the two carrier signals are 90° out of phase. The multiplexing system is shown in Figure P8.40(a) and the demultiplexing system in Figure P8.40(b). $x_1(t)$ and $x_2(t)$ are both assumed to be band limited with maximum frequency ω_M, so that $X_1(j\omega) = X_2(j\omega) = 0$ for $|\omega| > \omega_M$. The carrier frequency ω_c is assumed to be greater than ω_M. Show that $y_1(t) = x_1(t)$ and $y_2(t) = x_2(t)$.

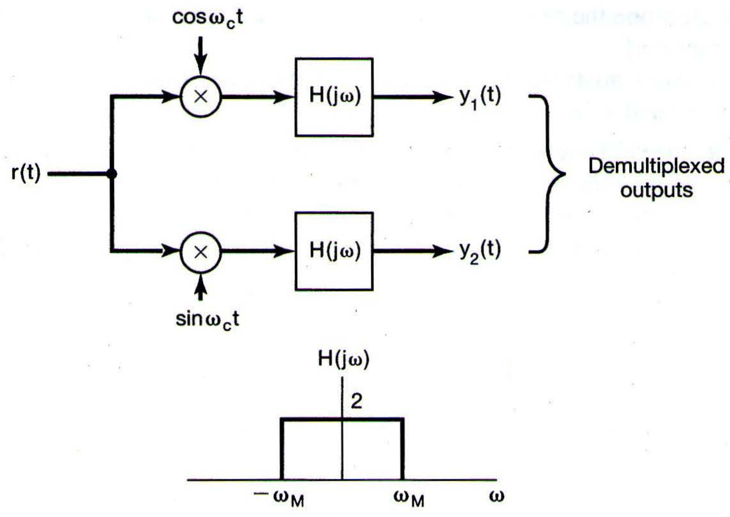


(a)

Figure P8.40

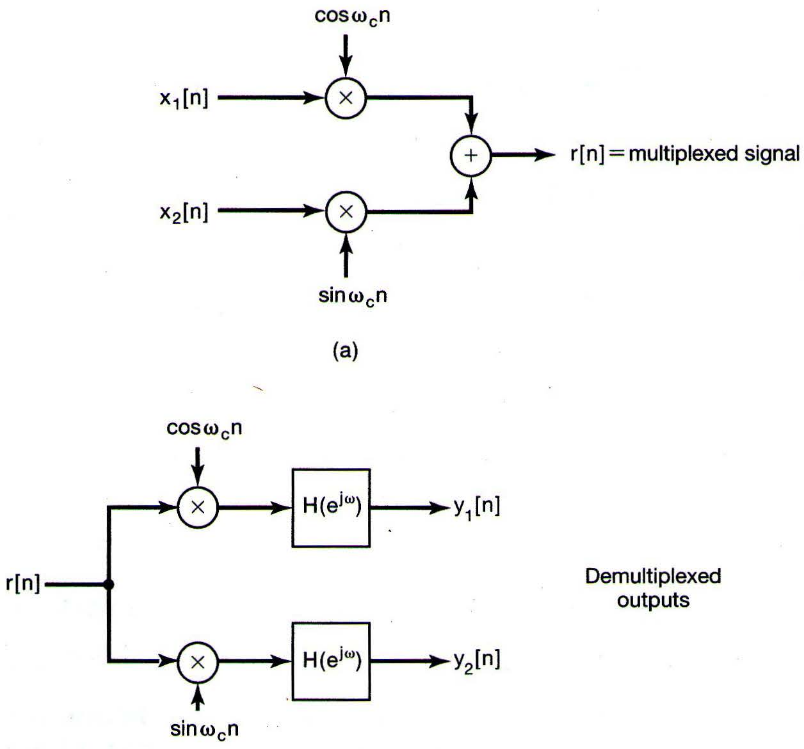
8.41. In Problem 8.40, we introduced the concept of quadrature multiplexing, whereby two signals are summed after each has been modulated with carrier signals of identical frequency, but with a phase difference of 90°. The corresponding discrete-time multiplexer and demultiplexer are shown in Figure P8.41. The signals $x_1[n]$ and $x_2[n]$ are both assumed to be band limited with maximum frequency ω_M, so that

$$X_1(e^{j\omega}) = X_2(e^{j\omega}) = 0 \text{ for } \omega_M < \omega < 2\pi - \omega_M.$$


$\cos\omega_c t$
$H(j\omega)$
$y_1(t)$
$r(t)$
Demultiplexed outputs
$H(j\omega)$
$y_2(t)$
$\sin\omega_c t$
$H(j\omega)$
2
$-\omega_M$
ω_M
ω

(b)

Figure P8.40 *Continued*


$\cos\omega_c n$
$x_1[n]$
$r[n]$ = multiplexed signal
$x_2[n]$
$\sin\omega_c n$
$\cos\omega_c n$
$H(e^{j\omega})$
$y_1[n]$
$r[n]$
Demultiplexed outputs
$H(e^{j\omega})$
$y_2[n]$
$\sin\omega_c n$

(a)

(b)

Figure P8.41

(a) Determine the range of values for ω_c so that $x_1[n]$ and $x_2[n]$ can be recovered from $r[n]$.

(b) With ω_c satisfying the conditions in part (a), determine $H(e^{j\omega})$ so that $y_1[n] = x_1[n]$ and $y_2[n] = x_2[n]$.

8.42. In order to avoid intersymbol interference, pulses used in PAM systems are designed to be zero at integer multiples of the symbol spacing T_1. In this problem, we develop a class of pulses which are zero at $t = kT_1$, $k = \pm 1, \pm 2, \pm 3, \ldots$.

Consider a pulse $p_1(t)$ that is real and even and that has a Fourier transform $P_1(j\omega)$. Also, assume that

$$P_1\left(-j\omega + j\frac{\pi}{T_1}\right) = -P_1\left(j\omega + j\frac{\pi}{T_1}\right), \quad 0 \le \omega \le \frac{\pi}{T_1}.$$

(a) Define a periodic sequence $\tilde{p}_1(t)$ with Fourier transform

$$\tilde{P}_1(j\omega) = \sum_{m=-\infty}^{\infty} P_1\left(j\omega - jm\frac{4\pi}{T_1}\right),$$

and show that

$$\tilde{P}_1(j\omega) = -\tilde{P}_1\left(j\omega - j\frac{2\pi}{T_1}\right).$$

(b) Use the result of the previous part to show that for some T

$$\tilde{p}_1(t) = 0, \; t = kT, \; k = 0, \pm 2, \pm 4, \ldots.$$

(c) Use the result of the previous part to show that

$$p_1(kT_1) = 0, \; k = \pm 1, \pm 2, \pm 3, \ldots.$$

(d) Show that a pulse $p(t)$ with Fourier transform

$$P(j\omega) = \begin{cases} 1 + P_1(j\omega), & |\omega| \le \frac{\pi}{T_1} \\ P_1(j\omega), & \frac{\pi}{T_1} \le |\omega| \le \frac{2\pi}{T_1} \\ 0, & \text{otherwise} \end{cases}$$

also has the property that

$$p(kT_1) = 0, \quad k = \pm 1, \pm 2, \pm 3, \ldots.$$

8.43. The impulse response of a channel used for PAM communication is specified by

$$h(t) = 10{,}000e^{-1{,}000t}u(t).$$

It is assumed that the phase response of the channel is approximately linear in the bandwidth of the channel. A pulse that is received after passing through the channel is processed using an LTI system S with impulse response $g(t)$ in order to compensate for the nonuniform gain over the channel bandwidth.

(a) Verify that if $g(t)$ has the Fourier transform

$$G(j\omega) = A + jB\omega,$$

where A and B are real constants, then $g(t)$ can compensate for the nonuniform gain over the channel bandwidth. Determine the values of A and B.

(b) It is proposed that S be implemented with the system shown in Figure P8.43. Determine the values of the gain factors α and β in this system.

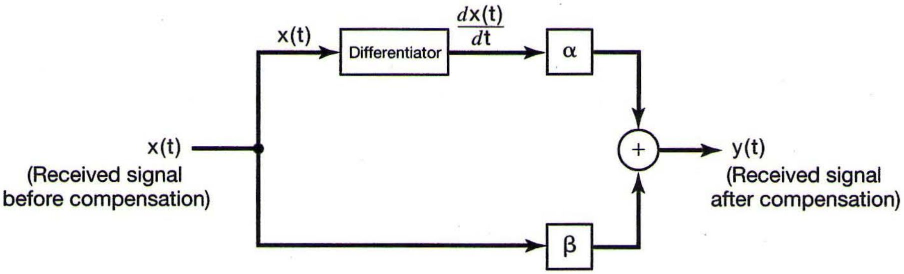


Figure P8.43

8.44. In this problem, we explore an *equalization* method used to avoid intersymbol interference caused in PAM systems by the channel having nonlinear phase over its bandwidth.

When a PAM pulse with zero-crossings at integer multiples of the symbol spacing T_1 is passed through a channel with nonlinear phase, the received pulse may no longer have zero-crossings at times that are integer multiples of T_1. Therefore, in order to avoid intersymbol interference, the received pulse is passed through a *zero-forcing equalizer,* which forces the pulse to have zero-crossings at integer multiples of T_1. This equalizer generates a new pulse $y(t)$ by summing up weighted and shifted versions of the received pulse $x(t)$. The pulse $y(t)$ is given by

$$y(t) = \sum_{l=-N}^{N} a_l x(t - lT_1), \tag{P8.44–1}$$

where the a_l are all real and are chosen such that

$$y(kT_1) = \begin{cases} 1, & k = 0 \\ 0, & k = \pm 1, \pm 2, \pm 3, \ldots, \pm N \end{cases}.$$

(a) Show that the equalizer is a filter and determine its impulse response.

(b) To illustrate the selection of the weights a_l, let us consider an example. If $x(0T_1) = 0.0$, $x(-T_1) = 0.2$, $x(T_1) = -0.2$, and $x(kT_1) = 0$ for $|k| > 1$, determine the values of a_0, a_1, and a_{-1} such that $y(\pm T_1) = 0$.

8.45. A band-limited signal $x(t)$ is to be transmitted using narrowband FM techniques. That is, the modulation index m, as defined in Section 8.7, is much less than $\pi/2$. Before $x(t)$ is transmitted to the modulator, it is processed so that $X(j\omega)|_{\omega=0} = 0$ and $|x(t)| < 1$. This normalized $x(t)$ is now used to angle-modulate a carrier to form the FM signal

$$y(t) = \cos\left(\omega_c t + m\int_{-\infty}^{t} x(\tau)\,d\tau\right).$$

(a) Determine the instantaneous frequency ω_i.

(b) Using eqs. (8.44) and (8.45), the narrowband assumption ($m << \pi/2$), and the preceding normalization conditions, show that

$$y(t) \approx \cos\omega_c t - \left(m\int_{-\infty}^{t} x(\tau)\,d\tau\right)\sin\omega_c t.$$

(c) What is the relationship among the bandwidth of $y(t)$, the bandwidth of $x(t)$, and the carrier frequency ω_c?

8.46. Consider the complex exponential function of time,

$$s(t) = e^{j\theta(t)}, \qquad \text{(P8.46–1)}$$

where $\theta(t) = \omega_0 t^2/2$.

Since the instantaneous frequency $\omega_i = d\theta/dt$ is also a function of time, the signal $s(t)$ may be regarded as an FM signal. In particular, since the signal sweeps linearly through the frequency spectrum with time, it is often called a frequency "chirp" or "chirp signal."

(a) Determine the instantaneous frequency.

(b) Determine and sketch the magnitude and phase of the Fourier transform of the "chirp signal." To evaluate the Fourier transform integral, you may find it helpful to complete the square in the exponent in the integrand and to use the relation

$$\int_{-\infty}^{+\infty} e^{jz^2}\,dz = \sqrt{\frac{\pi}{2}}(1+j).$$

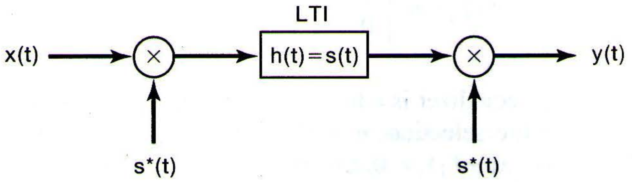


Figure P8.46

(c) Consider the system in Figure P8.46, in which $s(t)$ is the "chirp signal" in eq. (P8.46–1). Show that $y(t) = X(j\omega_0 t)$, where $X(j\omega)$ is the Fourier transform of $x(t)$.
(*Note:* The system in Figure P8.46 is referred to as the *"chirp" transform algorithm* and is often used in practice to obtain the Fourier transform of a signal.)

8.47. In Section 8.8 we considered synchronous discrete-time modulation and demodulation with a sinusoidal carrier. In this problem we want to consider the effect of a loss in synchronization in phase and/or frequency. The modulation and demodulation systems are shown in Figure P8.47(a), where both a phase and frequency difference between the modulator and demodulator carriers is indicated. Let the frequency difference $\omega_d - \omega_c$ be denoted as $\Delta\omega$ and the phase difference $\theta_d - \theta_c$ as $\Delta\theta$.

(a) If the spectrum of $x[n]$ is that shown in Figure P8.47(b), sketch the spectrum of $w[n]$, assuming $\Delta\omega = 0$.

(b) If $\Delta\omega = 0$, show that w can be chosen so that the output $r[n]$ is $r[n] = x[n]\cos\Delta\theta$. In particular, what is $r[n]$ if $\Delta\theta = \pi/2$?

(c) For $\Delta\theta = 0$, and $w = \omega_M + \Delta\omega$, show that the output $r[n] = x[n]\cos[\Delta\omega n]$ (assume that $\Delta\omega$ is small).

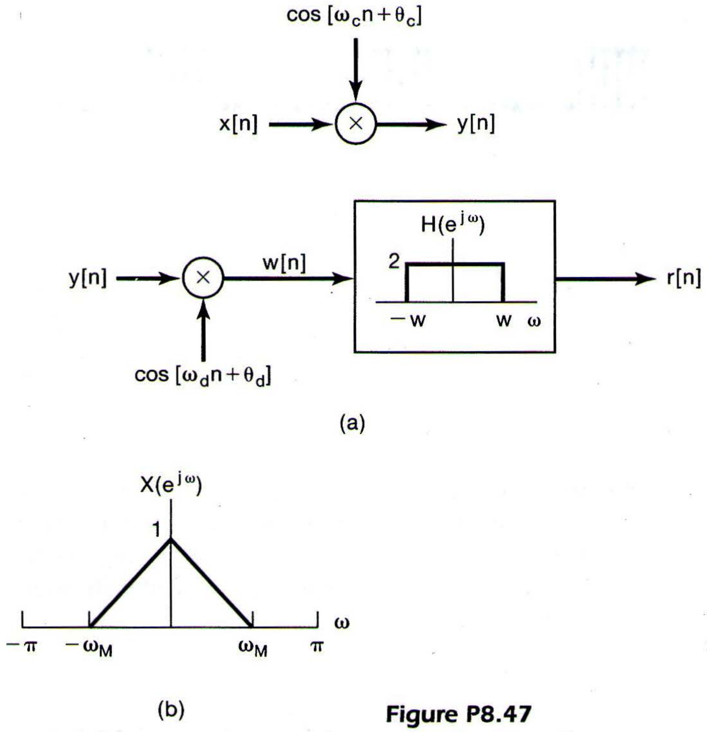


Figure P8.47

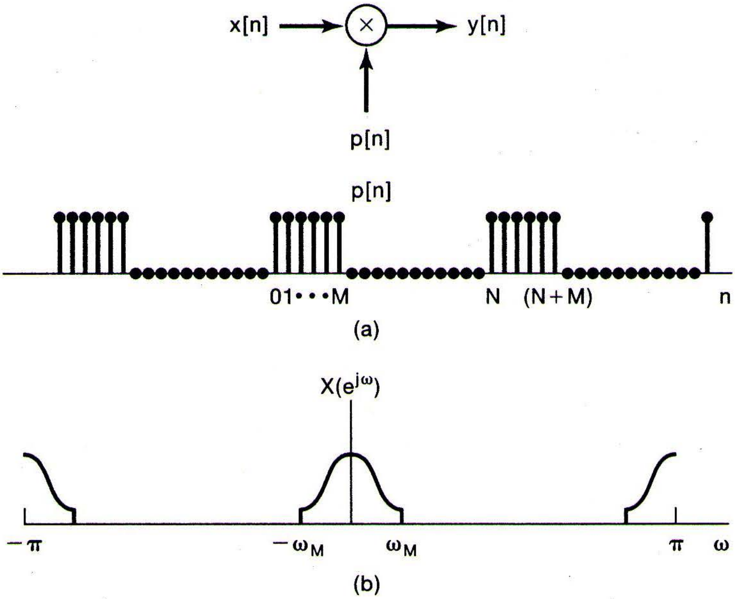
8.48. In this problem, we consider the analysis of discrete-time amplitude modulation of a pulse-train carrier. The system to be considered is shown in Figure P8.48(a).

(a) Determine and sketch the discrete-time Fourier transform of the periodic square-wave signal $p[n]$ in Figure P8.48(a).

(b) Assume that $x[n]$ has the spectrum shown in Figure P8.48(b). With $\omega_M = \pi/2N$ and with $M = 1$ in Figure P8.48(a), sketch $Y(e^{j\omega})$, the Fourier transform of $y[n]$.

(c) Now assume that $X(e^{j\omega})$ is known to be band limited with $X(e^{j\omega}) = 0$, $\omega_M < \omega < 2\pi - \omega_M$, but is otherwise unspecified. For the system of Figure P8.48(a), determine, as a function of N, the maximum allowable value of ω_M that will permit $x[n]$ to be recovered from $y[n]$. Indicate whether your result depends on M.

(d) With ω_M and N satisfying the condition determined in part (c), state or show in block diagram form how to recover $x[n]$ from $y[n]$.



(a)

(b)

Figure P8.48

8.49. In practice it is often very difficult to build an amplifier at very low frequencies. Consequently, low-frequency amplifiers typically exploit the principles of amplitude modulation to shift the signal into a higher-frequency band. Such an amplifier is referred to as a *chopper amplifier* and is illustrated in the block-diagram form in Figure P8.49.

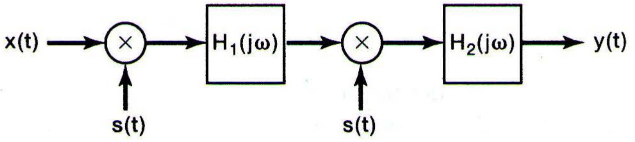


Figure P8.49

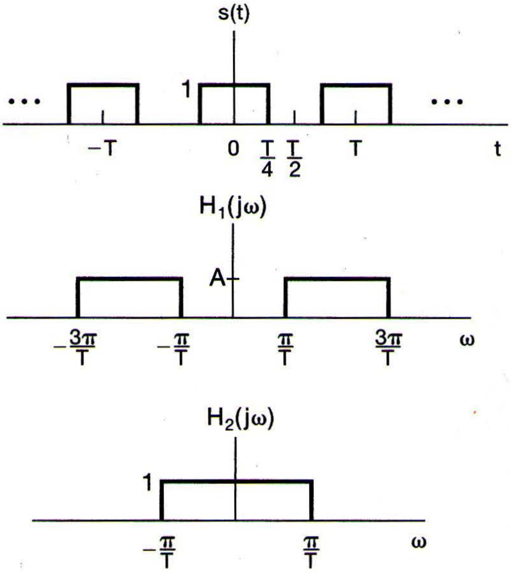

s(t)
1
• • •
• • •
−T
0
T/4
T/2
T
t
H₁(jω)
A
−3π/T
−π/T
π/T
3π/T
ω
H₂(jω)
1
−π/T
π/T
ω

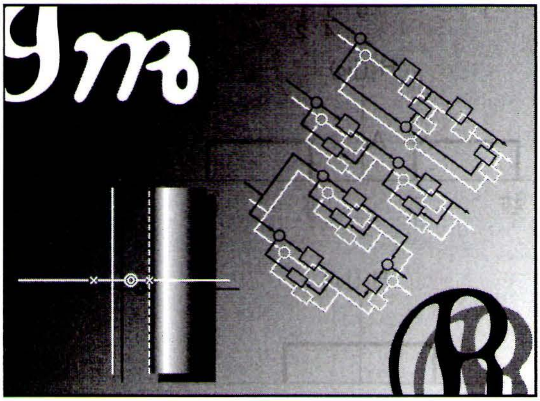
Figure P8.49 *Continued*

(a) Determine in terms of T the highest allowable frequency present in $x(t)$, if $y(t)$ is to be proportional to $x(t)$ (i.e., if the overall system is to be equivalent to an amplifier).

(b) With $x(t)$ bandlimited as specified in part (a), determine the gain of the overall system in Figure P8.49 in terms of A and T.

9

THE LAPLACE TRANSFORM



9.0 INTRODUCTION

In the preceding chapters, we have seen that the tools of Fourier analysis are extremely useful in the study of many problems of practical importance involving signals and LTI systems. This is due in large part to the fact that broad classes of signals can be represented as linear combinations of periodic complex exponentials and that complex exponentials are eigenfunctions of LTI systems. The continuous-time Fourier transform provides us with a representation for signals as linear combinations of complex exponentials of the form e^{st} with $s = j\omega$. However the eigenfunction property introduced in Section 3.2 and many of its consequences apply as well for arbitrary values of s and not only those values that are purely imaginary. This observation leads to a generalization of the continuous-time Fourier transform, known as the Laplace transform, which we develop in this chapter. In the next chapter we develop the corresponding discrete-time generalization known as the z-transform.

As we will see, the Laplace and z-transforms have many of the properties that make Fourier analysis so useful. Moreover, not only do these transforms provide additional tools and insights for signals and systems that can be analyzed using the Fourier transform, but they also can be applied in some very important contexts in which Fourier transforms cannot. For example Laplace and z-transforms can be applied to the analysis of many unstable systems and consequently play an important role in the investigation of the stability or instability of systems. This fact, combined with the algebraic properties that Laplace and z-transforms share with Fourier transforms, leads to a very important set of tools for system analysis and in particular for the analysis of feedback systems, which we develop in Chapter 11.

9.1 THE LAPLACE TRANSFORM

In Chapter 3, we saw that the response of a linear time-invariant system with impulse response $h(t)$ to a complex exponential input of the form e^{st} is

$$y(t) = H(s)e^{st}, \tag{9.1}$$

where

$$H(s) = \int_{-\infty}^{\infty} h(t)e^{-st}\,dt. \tag{9.2}$$

For s imaginary (i.e., $s = j\omega$), the integral in eq. (9.2) corresponds to the Fourier transform of $h(t)$. For general values of the complex variable s, it is referred to as the *Laplace transform* of the impulse response $h(t)$.

The Laplace transform of a general signal $x(t)$ is defined as[1]

$$\boxed{X(s) \triangleq \int_{-\infty}^{+\infty} x(t)e^{-st}\,dt,} \tag{9.3}$$

and we note in particular that it is a function of the independent variable s corresponding to the complex variable in the exponent of e^{-st}. The complex variable s can be written as $s = \sigma + j\omega$, with σ and ω the real and imaginary parts, respectively. For convenience, we will sometimes denote the Laplace transform in operator form as $\mathcal{L}\{x(t)\}$ and denote the transform relationship between $x(t)$ and $X(s)$ as

$$x(t) \overset{\mathcal{L}}{\longleftrightarrow} X(s). \tag{9.4}$$

When $s = j\omega$, eq. (9.3) becomes

$$X(j\omega) = \int_{-\infty}^{+\infty} x(t)e^{-j\omega t}\,dt, \tag{9.5}$$

which corresponds to the *Fourier transform* of $x(t)$; that is,

$$X(s)|_{s=j\omega} = \mathcal{F}\{x(t)\}. \tag{9.6}$$

The Laplace transform also bears a straightforward relationship to the Fourier transform when the complex variable s is not purely imaginary. To see this relationship, consider $X(s)$ as specified in eq. (9.3) with s expressed as $s = \sigma + j\omega$, so that

$$X(\sigma + j\omega) = \int_{-\infty}^{+\infty} x(t)e^{-(\sigma + j\omega)t}\,dt, \tag{9.7}$$

---

[1] The transform defined by eq. (9.3) is often called the *bilateral Laplace transform,* to distinguish it from the unilateral Laplace transform, which we discuss in Section 9.9. The bilateral transform in eq. (9.3) involves an integration from $-\infty$ to $+\infty$, while the unilateral transform has a form similar to that in eq. (9.3), but with limits of integration from 0 to $+\infty$. As we are primarily concerned with the bilateral transform, we will omit the word "bilateral," except where it is needed in Section 9.9 to avoid ambiguity.

or

$$X(\sigma + j\omega) = \int_{-\infty}^{+\infty} [x(t)e^{-\sigma t}]e^{-j\omega t}\,dt. \tag{9.8}$$

We recognize the right-hand side of eq. (9.8) as the Fourier transform of $x(t)e^{-\sigma t}$; that is, the Laplace transform of $x(t)$ can be interpreted as the Fourier transform of $x(t)$ after multiplication by a real exponential signal. The real exponential $e^{-\sigma t}$ may be decaying or growing in time, depending on whether σ is positive or negative.

To illustrate the Laplace transform and its relationship to the Fourier transform, let us consider the following example:

Example 9.1

Let the signal $x(t) = e^{-at}u(t)$. From Example 4.1, the Fourier transform $X(j\omega)$ converges for $a > 0$ and is given by

$$X(j\omega) = \int_{-\infty}^{+\infty} e^{-at}u(t)e^{-j\omega t}\,dt = \int_{0}^{\infty} e^{-at}e^{-j\omega t}\,dt = \frac{1}{j\omega + a}, \quad a > 0. \tag{9.9}$$

From eq. (9.3), the Laplace transform is

$$X(s) = \int_{-\infty}^{\infty} e^{-at}u(t)e^{-st}dt = \int_{0}^{\infty} e^{-(s+a)t}dt, \tag{9.10}$$

or, with $s = \sigma + j\omega$,

$$X(\sigma + j\omega) = \int_{0}^{\infty} e^{-(\sigma+a)t}e^{-j\omega t}\,dt. \tag{9.11}$$

By comparison with eq. (9.9) we recognize eq. (9.11) as the Fourier transform of $e^{-(\sigma+a)t}u(t)$, and thus,

$$X(\sigma + j\omega) = \frac{1}{(\sigma + a) + j\omega}, \quad \sigma + a > 0, \tag{9.12}$$

or equivalently, since $s = \sigma + j\omega$ and $\sigma = \mathcal{R}e\{s\}$,

$$X(s) = \frac{1}{s + a}, \quad \mathcal{R}e\{s\} > -a. \tag{9.13}$$

That is,

$$e^{-at}u(t) \overset{\mathcal{L}}{\longleftrightarrow} \frac{1}{s + a}, \quad \mathcal{R}e\{s\} > -a. \tag{9.14}$$

For example, for $a = 0$, $x(t)$ is the unit step with Laplace transform $X(s) = 1/s$, $\mathcal{R}e\{s\} > 0$.

We note, in particular, that just as the Fourier transform does not converge for all signals, the Laplace transform may converge for some values of $\mathcal{R}e\{s\}$ and not for others. In eq. (9.13), the Laplace transform converges only for $\sigma = \mathcal{R}e\{s\} > -a$. If a is positive, then $X(s)$ can be evaluated at $\sigma = 0$ to obtain

$$X(0 + j\omega) = \frac{1}{j\omega + a}. \tag{9.15}$$

As indicated in eq. (9.6), for $\sigma = 0$ the Laplace transform is equal to the Fourier transform, as is evident in the preceding example by comparing eqs. (9.9) and (9.15). If a is negative or zero, the Laplace transform still exists, but the Fourier transform does not.

Example 9.2

For comparison with Example 9.1, let us consider as a second example the signal

$$x(t) = -e^{-at}u(-t). \tag{9.16}$$

Then

$$\begin{aligned} X(s) &= -\int_{-\infty}^{\infty} e^{-at}e^{-st}u(-t)\,dt \\ &= -\int_{-\infty}^{0} e^{-(s+a)t}\,dt, \end{aligned} \tag{9.17}$$

or

$$X(s) = \frac{1}{s+a}. \tag{9.18}$$

For convergence in this example, we require that $\mathcal{R}e\{s+a\} < 0$, or $\mathcal{R}e\{s\} < -a$; that is,

$$-e^{-at}u(-t) \overset{\mathcal{L}}{\longleftrightarrow} \frac{1}{s+a}, \qquad \mathcal{R}e\{s\} < -a. \tag{9.19}$$

Comparing eqs. (9.14) and (9.19), we see that the algebraic expression for the Laplace transform is identical for both of the signals considered in Examples 9.1 and 9.2. However, from the same equations, we also see that the set of values of s for which the expression is valid is very different in the two examples. This serves to illustrate the fact that, in specifying the Laplace transform of a signal, both the algebraic expression and the range of values of s for which this expression is valid are required. In general, the range of values of s for which the integral in eq.(9.3) converges is referred to as the *region of convergence* (which we abbreviate as ROC) of the Laplace transform. That is, the ROC consists of those values of $s = \sigma + j\omega$ for which the Fourier transform of $x(t)e^{-\sigma t}$ converges. We will have more to say about the ROC as we develop some insight into the properties of the Laplace transform.

A convenient way to display the ROC is shown in Figure 9.1. The variable s is a complex number, and in the figure we display the complex plane, generally referred to as the s-plane, associated with this complex variable. The coordinate axes are $\mathcal{R}e\{s\}$ along the horizontal axis and $\mathcal{I}m\{s\}$ along the vertical axis. The horizontal and vertical axes are sometimes referred to as the σ-axis and the $j\omega$-axis, respectively. The shaded region in Figure 9.1(a) represents the set of points in the s-plane corresponding to the region of convergence for Example 9.1. The shaded region in Figure 9.1(b) indicates the region of convergence for Example 9.2.

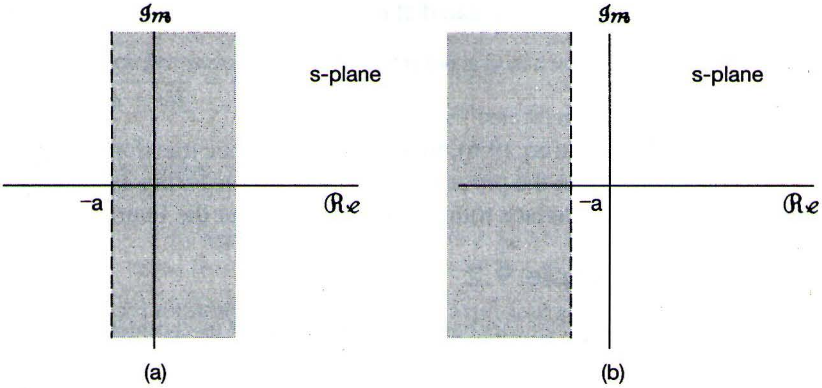


Figure 9.1 (a) ROC for Example 9.1; (b) ROC for Example 9.2.

Example 9.3

In this example, we consider a signal that is the sum of two real exponentials:

$$x(t) = 3e^{-2t}u(t) - 2e^{-t}u(t). \tag{9.20}$$

The algebraic expression for the Laplace transform is then

$$\begin{aligned} X(s) &= \int_{-\infty}^{\infty}\left[3e^{-2t}u(t) - 2e^{-t}u(t)\right]e^{-st}\,dt \\ &= 3\int_{-\infty}^{\infty} e^{-2t}e^{-st}u(t)\,dt - 2\int_{-\infty}^{\infty} e^{-t}e^{-st}u(t)\,dt. \end{aligned} \tag{9.21}$$

Each of the integrals in eq. (9.21) is of the same form as the integral in eq. (9.10), and consequently, we can use the result in Example 9.1 to obtain

$$X(s) = \frac{3}{s+2} - \frac{2}{s+1}. \tag{9.22}$$

To determine the ROC we note that $x(t)$ is a sum of two real exponentials, and from eq. (9.21) we see that $X(s)$ is the sum of the Laplace transforms of each of the individual terms. The first term is the Laplace transform of $3e^{-2t}u(t)$ and the second term the Laplace transform of $-2e^{-t}u(t)$. From Example 9.1, we know that

$$e^{-t}u(t) \overset{\mathcal{L}}{\longleftrightarrow} \frac{1}{s+1}, \qquad \mathcal{Re}\{s\} > -1,$$

$$e^{-2t}u(t) \overset{\mathcal{L}}{\longleftrightarrow} \frac{1}{s+2}, \qquad \mathcal{Re}\{s\} > -2.$$

The set of values of $\mathcal{Re}\{s\}$ for which the Laplace transforms of both terms converge is $\mathcal{Re}\{s\} > -1$, and thus, combining the two terms on the right-hand side of eq. (9.22), we obtain

$$3e^{-2t}u(t) - 2e^{-t}u(t) \overset{\mathcal{L}}{\longleftrightarrow} \frac{s-1}{s^2+3s+2}, \quad \mathcal{Re}\{s\} > -1. \tag{9.23}$$

Example 9.4

In this example, we consider a signal that is the sum of a real and a complex exponential:

$$x(t) = e^{-2t}u(t) + e^{-t}(\cos 3t)u(t). \tag{9.24}$$

Using Euler's relation, we can write

$$x(t) = \left[e^{-2t} + \frac{1}{2}e^{-(1-3j)t} + \frac{1}{2}e^{-(1+3j)t}\right]u(t),$$

and the Laplace transform of $x(t)$ then can be expressed as

$$\begin{aligned} X(s) = &\int_{-\infty}^{\infty} e^{-2t}u(t)e^{-st}\,dt \\ &+ \frac{1}{2}\int_{-\infty}^{\infty} e^{-(1-3j)t}u(t)e^{-st}\,dt \\ &+ \frac{1}{2}\int_{-\infty}^{\infty} e^{-(1+3j)t}u(t)e^{-st}\,dt. \end{aligned} \tag{9.25}$$

Each of the integrals in eq. (9.25) represents a Laplace transform of the type encountered in Example 9.1. It follows that

$$e^{-2t}u(t) \overset{\mathcal{L}}{\longleftrightarrow} \frac{1}{s+2}, \quad \mathcal{Re}\{s\} > -2, \tag{9.26}$$

$$e^{-(1-3j)t}u(t) \overset{\mathcal{L}}{\longleftrightarrow} \frac{1}{s+(1-3j)}, \quad \mathcal{Re}\{s\} > -1, \tag{9.27}$$

$$e^{-(1+3j)t}u(t) \overset{\mathcal{L}}{\longleftrightarrow} \frac{1}{s+(1+3j)}, \quad \mathcal{Re}\{s\} > -1. \tag{9.28}$$

For all three Laplace transforms to converge simultaneously, we must have $\mathcal{Re}\{s\} > -1$. Consequently, the Laplace transform of $x(t)$ is

$$\frac{1}{s+2} + \frac{1}{2}\left(\frac{1}{s+(1-3j)}\right) + \frac{1}{2}\left(\frac{1}{s+(1+3j)}\right), \quad \mathcal{Re}\{s\} > -1, \tag{9.29}$$

or, with terms combined over a common denominator,

$$e^{-2t}u(t) + e^{-t}(\cos 3t)u(t) \overset{\mathcal{L}}{\longleftrightarrow} \frac{2s^2+5s+12}{(s^2+2s+10)(s+2)}, \quad \mathcal{Re}\{s\} > -1. \tag{9.30}$$

In each of the four preceding examples, the Laplace transform is rational, i.e., it is a ratio of polynomials in the complex variable s, so that

$$X(s) = \frac{N(s)}{D(s)}, \tag{9.31}$$

where $N(s)$ and $D(s)$ are the numerator polynomial and denominator polynomial, respectively. As suggested by Examples 9.3 and 9.4, $X(s)$ will be rational whenever $x(t)$ is a linear combination of real or complex exponentials. As we will see in Section 9.7, rational transforms also arise when we consider LTI systems specified in terms of linear constant-coefficient differential equations. Except for a scale factor, the numerator and denominator polynomials in a rational Laplace transform can be specified by their roots; thus, marking the locations of the roots of $N(s)$ and $D(s)$ in the s-plane and indicating the ROC provides a convenient pictorial way of describing the Laplace transform. For example, in Figure 9.2(a) we show the s-plane representation of the Laplace transform of Example 9.3, with the location of each root of the denominator polynomial in eq. (9.23) indicated with "×" and the location of the root of the numerator polynomial in eq. (9.23) indicated with "o." The corresponding plot of the roots of the numerator and denominator polynomials for the Laplace transform in Example 9.4 is given in Figure 9.2(b). The region of convergence for each of these examples is shaded in the corresponding plot.

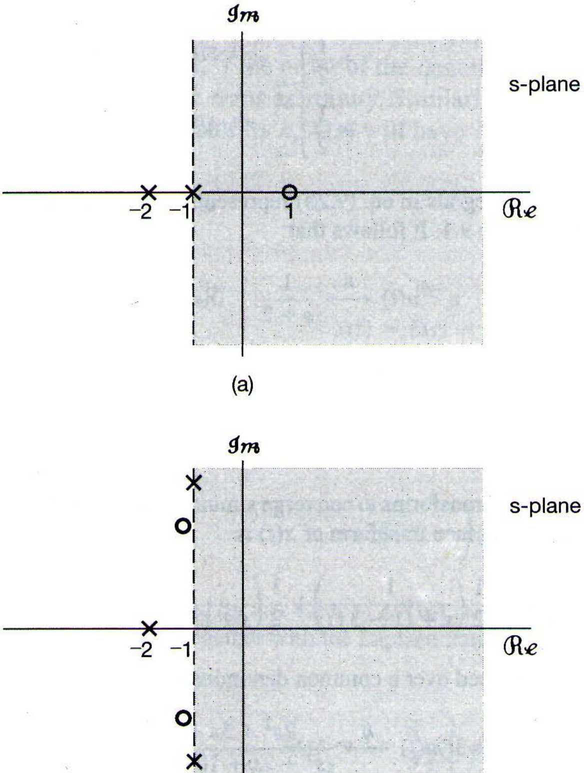


Figure 9.2 s-plane representation of the Laplace transforms for (a) Example 9.3 and (b) Example 9.4. Each × in these figures marks the location of a pole of the corresponding Laplace transform—i.e., a root of the denominator. Similarly, each o marks a zero—i.e., a root of the the numerator. The shaded regions indicate the ROCs.

For rational Laplace transforms, the roots of the numerator polynomial are commonly referred to as the *zeros* of $X(s)$, since, for those values of s, $X(s) = 0$. The roots of the denominator polynomial are referred to as the *poles* of $X(s)$, and for those values of s, $X(s)$ is infinite. The poles and zeros of $X(s)$ in the finite s-plane completely characterize the algebraic expression for $X(s)$ to within a scale factor. The representation of $X(s)$ through its poles and zeros in the s-plane is referred to as the *pole-zero plot* of $X(s)$.

However, as we saw in Examples 9.1 and 9.2, knowledge of the algebraic form of $X(s)$ does not by itself identify the ROC for the Laplace transform. That is, a complete specification, to within a scale factor, of a rational Laplace transform consists of the pole-zero plot of the transform, together with its ROC (which is commonly shown as a shaded region in the s-plane, as in Figures 9.1 and 9.2).

Also, while they are not needed to specify the algebraic form of a rational transform $X(s)$, it is sometimes convenient to refer to poles or zeros of $X(s)$ at infinity. Specifically, if the order of the denominator polynomial is greater than the order of the numerator polynomial, then $X(s)$ will become zero as s approaches infinity. Conversely, if the order of the numerator polynomial is greater than the order of the denominator, then $X(s)$ will become unbounded as s approaches infinity. This behavior can be interpreted as zeros or poles at infinity. For example, the Laplace transform in eq. (9.23) has a denominator of order 2 and a numerator of order only 1, so in this case $X(s)$ has one zero at infinity. The same is true for the transform in eq. (9.30), in which the numerator is of order 2 and the denominator is of order 3. In general, if the order of the denominator exceeds the order of the numerator by k, $X(s)$ will have k zeros at infinity. Similarly, if the order of the numerator exceeds the order of the denominator by k, $X(s)$ will have k poles at infinity.

Example 9.5

Let

$$x(t) = \delta(t) - \frac{4}{3}e^{-t}u(t) + \frac{1}{3}e^{2t}u(t). \tag{9.32}$$

The Laplace transform of the second and third terms on the right-hand side of eq. (9.32) can be evaluated from Example 9.1. The Laplace transform of the unit impulse can be evaluated directly as

$$\mathcal{L}\{\delta(t)\} = \int_{-\infty}^{+\infty} \delta(t)e^{-st}\,dt = 1, \tag{9.33}$$

which is valid for any value of s. That is, the ROC of $\mathcal{L}\{\delta(t)\}$ is the entire s-plane. Using this result, together with the Laplace transforms of the other two terms in eq. (9.32), we obtain

$$X(s) = 1 - \frac{4}{3}\frac{1}{s+1} + \frac{1}{3}\frac{1}{s-2}, \quad \Re e\{s\} > 2, \tag{9.34}$$

or

$$X(s) = \frac{(s-1)^2}{(s+1)(s-2)}, \quad \Re e\{s\} > 2, \tag{9.35}$$

where the ROC is the set of values of s for which the Laplace transforms of all three terms in $x(t)$ converge. The pole-zero plot for this example is shown in Figure 9.3, together with the ROC. Also, since the degrees of the numerator and denominator of $X(s)$ are equal, $X(s)$ has neither poles nor zeros at infinity.

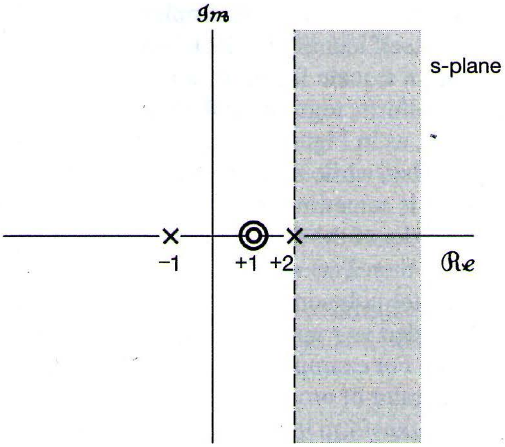


Figure 9.3 Pole-zero plot and ROC for Example 9.5.

Recall from eq. (9.6) that, for $s = j\omega$, the Laplace transform corresponds to the Fourier transform. However, if the ROC of the Laplace transform does not include the $j\omega$-axis, (i.e., if $\mathcal{R}e\{s\} = 0$), then the Fourier transform does not converge. As we see from Figure 9.3, this, in fact, is the case for Example 9.5, which is consistent with the fact that the term $(1/3)e^{2t}u(t)$ in $x(t)$ does not have a Fourier transform. Note also in this example that the two zeros in eq. (9.35) occur at the same value of s. In general, we will refer to the *order* of a pole or zero as the number of times it is repeated at a given location. In Example 9.5 there is a second-order zero at $s = 1$ and two first-order poles, one at $s = -1$, the other at $s = 2$. In this example the ROC lies to the right of the rightmost pole. In general, for rational Laplace transforms, there is a close relationship between the locations of the poles and the possible ROCs that can be associated with a given pole-zero plot. Specific constraints on the ROC are closely associated with time-domain properties of $x(t)$. In the next section, we explore some of these constraints and properties.

9.2 THE REGION OF CONVERGENCE FOR LAPLACE TRANSFORMS

In the preceding section, we saw that a complete specification of the Laplace transform requires not only the algebraic expression for $X(s)$, but also the associated region of convergence. As evidenced by Examples 9.1 and 9.2, two very different signals can have identical algebraic expressions for $X(s)$, so that their Laplace transforms are distinguishable *only* by the region of convergence. In this section, we explore some specific constraints on the ROC for various classes of signals. As we will see, an understanding of these constraints often permits us to specify implicitly or to reconstruct the ROC from knowledge of only the algebraic expression for $X(s)$ and certain general characteristics of $x(t)$ in the time domain.

Property 1: The ROC of $X(s)$ consists of strips parallel to the $j\omega$-axis in the s-plane.

The validity of this property stems from the fact that the ROC of $X(s)$ consists of those values of $s = \sigma + j\omega$ for which the Fourier transform of $x(t)e^{-\sigma t}$ converges. That is, the ROC of the Laplace transform of $x(t)$ consists of those values of s for which $x(t)e^{-\sigma t}$ is absolutely integrable:[2]

$$\int_{-\infty}^{+\infty} |x(t)|e^{-\sigma t}\,dt < \infty. \tag{9.36}$$

Property 1 then follows, since this condition depends only on σ, the real part of s.

Property 2: For rational Laplace transforms, the ROC does not contain any poles.

Property 2 is easily observed in all the examples studied thus far. Since $X(s)$ is infinite at a pole, the integral in eq. (9.3) clearly does not converge at a pole, and thus the ROC cannot contain values of s that are poles.

Property 3: If $x(t)$ is of finite duration and is absolutely integrable, then the ROC is the entire s-plane.

The intuition behind this result is suggested in Figures 9.4 and 9.5. Specifically, a finite-duration signal has the property that it is zero outside an interval of finite duration, as illustrated in Figure 9.4. In Figure 9.5(a), we have shown $x(t)$ of Figure 9.4 multiplied by a decaying exponential, and in Figure 9.5(b) the same signal multiplied by a growing

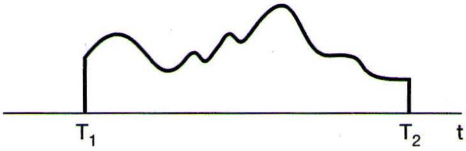
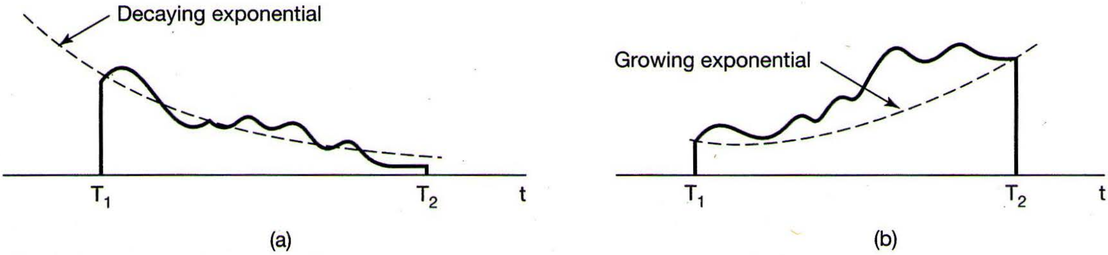


Figure 9.4 Finite-duration signal.



(a)

(b)

Figure 9.5 (a) Finite-duration signal of Figure 9.4 multiplied by a decaying exponential; (b) finite-duration signal of Figure 9.4 multiplied by a growing exponential.

---

[2] For a more thorough and formal treatment of Laplace transforms and their mathematical properties, including convergence, see E. D. Rainville, *The Laplace Transform: An Introduction* (New York: Macmillan, 1963), and R. V. Churchill and J. W. Brown, *Complex Variables and Applications* (5th ed.) (New York: McGraw-Hill, 1990). Note that the condition of absolute integrability is one of the Dirichlet conditions introduced in Section 4.1 in the context of our discussion of the convergence of Fourier transforms.

exponential. Since the interval over which $x(t)$ is nonzero is finite, the exponential weighting is never unbounded, and consequently, it is reasonable that the integrability of $x(t)$ not be destroyed by this exponential weighting.

A more formal verification of Property 3 is as follows: Suppose that $x(t)$ is absolutely integrable, so that

$$\int_{T_1}^{T_2} |x(t)|\,dt < \infty. \tag{9.37}$$

For $s = \sigma + j\omega$ to be in the ROC, we require that $x(t)e^{-\sigma t}$ be absolutely integrable, i.e.,

$$\int_{T_1}^{T_2} |x(t)|e^{-\sigma t}\,dt < \infty. \tag{9.38}$$

Eq. (9.37) verifies that s is in the ROC when $\mathcal{R}e\{s\} = \sigma = 0$. For $\sigma > 0$, the maximum value of $e^{-\sigma t}$ over the interval on which $x(t)$ is nonzero is $e^{-\sigma T_1}$, and thus we can write

$$\int_{T_1}^{T_2} |x(t)|e^{-\sigma t}\,dt < e^{-\sigma T_1}\int_{T_1}^{T_2} |x(t)|\,dt. \tag{9.39}$$

Since the right-hand side of eq.(9.39) is bounded, so is the left-hand side; therefore, the s-plane for $\mathcal{R}e\{s\} > 0$ must also be in the ROC. By a similar argument, if $\sigma < 0$, then

$$\int_{T_1}^{T_2} |x(t)|e^{-\sigma t}\,dt < e^{-\sigma T_2}\int_{T_1}^{T_2} |x(t)|\,dt, \tag{9.40}$$

and again, $x(t)e^{-\sigma t}$ is absolutely integrable. Thus, the ROC includes the entire s-plane.

Example 9.6

Let

$$x(t) = \begin{cases} e^{-at}, & 0 < t < T \\ 0, & \text{otherwise} \end{cases}. \tag{9.41}$$

Then

$$X(s) = \int_0^T e^{-at}e^{-st}\,dt = \frac{1}{s+a}[1 - e^{-(s+a)T}]. \tag{9.42}$$

Since in this example $x(t)$ is of finite length, it follows from Property 3 that the ROC is the entire s-plane. In the form of eq. (9.42), $X(s)$ would appear to have a pole at $s = -a$, which, from Property 2, would be inconsistent with an ROC that consists of the entire s-plane. In fact, however, in the algebraic expression in eq. (9.42), both numerator and denominator are zero at $s = -a$, and thus, to determine $X(s)$ at $s = -a$, we can use L'hôpital's rule to obtain

$$\lim_{s \to -a} X(s) = \lim_{s \to -a}\left[\frac{\frac{d}{ds}(1 - e^{-(s+a)T})}{\frac{d}{ds}(s+a)}\right] = \lim_{s \to -a} Te^{-aT}e^{-sT},$$

so that

$$X(-a) = T. \tag{9.43}$$

It is important to recognize that, to ensure that the exponential weighting is bounded over the interval in which $x(t)$ is nonzero, the preceding discussion relies heavily on the fact that $x(t)$ is of finite duration. In the next two properties, we consider modifications of the result in Property 3 when $x(t)$ is of finite extent in only the positive-time or negative-time direction.

Property 4: If $x(t)$ is right sided, and if the line $\mathcal{R}e\{s\} = \sigma_0$ is in the ROC, then all values of s for which $\mathcal{R}e\{s\} > \sigma_0$ will also be in the ROC.

A *right-sided* signal is a signal for which $x(t) = 0$ prior to some finite time T_1, as illustrated in Figure 9.6. It is possible that, for such a signal, there is no value of s for which the Laplace transform will converge. One example is the signal $x(t) = e^{t^2}u(t)$. However, suppose that the Laplace transform converges for some value of σ, which we denote by σ_0. Then

$$\int_{-\infty}^{+\infty} |x(t)|e^{-\sigma_0 t}\, dt < \infty, \tag{9.44}$$

or equivalently, since $x(t)$ is right sided,

$$\int_{T_1}^{+\infty} |x(t)|e^{-\sigma_0 t}\, dt < \infty. \tag{9.45}$$

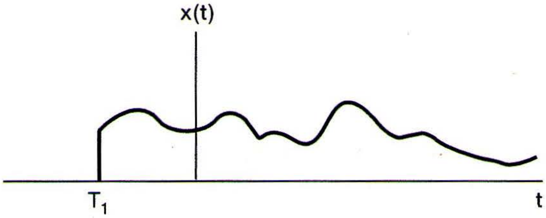


Figure 9.6 Right-sided signal.

Then if $\sigma_1 > \sigma_0$, it must also be true that $x(t)e^{-\sigma_1 t}$ is absolutely integrable, since $e^{-\sigma_1 t}$ decays faster than $e^{-\sigma_0 t}$ as $t \longrightarrow +\infty$, as illustrated in Figure 9.7. Formally, we can say that with $\sigma_1 > \sigma_0$,

$$\begin{aligned} \int_{T_1}^{\infty} |x(t)|e^{-\sigma_1 t}\, dt &= \int_{T_1}^{\infty} |x(t)|e^{-\sigma_0 t}e^{-(\sigma_1-\sigma_0)t}\, dt \\ &\leq e^{-(\sigma_1-\sigma_0)T_1}\int_{T_1}^{\infty} |x(t)|e^{-\sigma_0 t}\, dt. \end{aligned} \tag{9.46}$$

Since T_1 is finite, it follows from eq. (9.45) that the right side of the inequality in eq. (9.46) is finite, and hence, $x(t)e^{-\sigma_1 t}$ is absolutely integrable.

Note that in the preceding argument we explicitly rely on the fact that $x(t)$ is right sided, so that, although with $\sigma_1 > \sigma_0$, $e^{-\sigma_1 t}$ diverges faster than $e^{-\sigma_0 t}$ as $t \rightarrow -\infty$, $x(t)e^{-\sigma_1 t}$ cannot grow without bound in the negative-time direction, since $x(t) = 0$ for

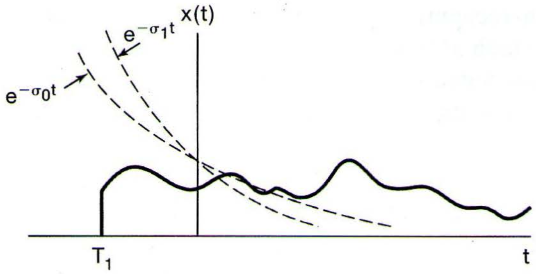


Figure 9.7 If $x(t)$ is right sided and $x(t)e^{-\sigma_0 t}$ is absolutely integrable, then $x(t)e^{-\sigma_1 t}$, $\sigma_1 > \sigma_0$, will also be absolutely integrable.

$t < T_1$. Also, in this case, if a point s is in the ROC, then all the points to the right of s, i.e., all points with larger real parts, are in the ROC. For this reason, the ROC in this case is commonly referred to as a *right-half plane.*

Property 5: If $x(t)$ is left sided, and if the line $\mathcal{R}e\{s\} = \sigma_0$ is in the ROC, then all values of s for which $\mathcal{R}e\{s\} < \sigma_0$ will also be in the ROC.

A *left-sided* signal is a signal for which $x(t) = 0$ after some finite time T_2, as illustrated in Figure 9.8. The argument and intuition behind this property are exactly analogous to the argument and intuition behind Property 4. Also, for a left-sided signal, the ROC is commonly referred to as a *left-half plane,* as if a point s is in the ROC, then all points to the left of s are in the ROC.

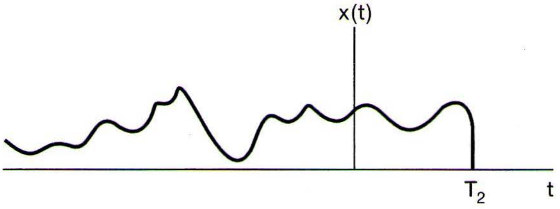
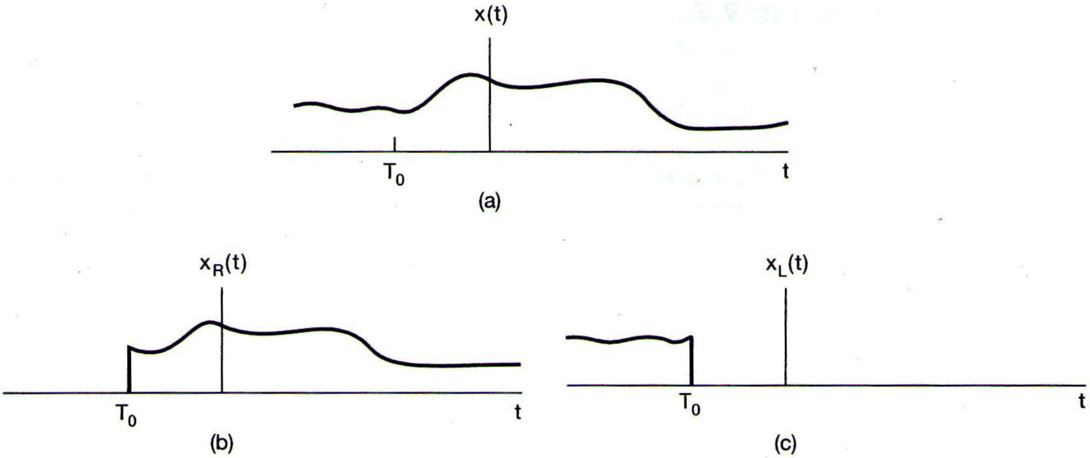


Figure 9.8 Left-sided signal.

Property 6: If $x(t)$ is two sided, and if the line $\mathcal{R}e\{s\} = \sigma_0$ is in the ROC, then the ROC will consist of a strip in the s-plane that includes the line $\mathcal{R}e\{s\} = \sigma_0$.

A *two-sided* signal is a signal that is of infinite extent for both $t > 0$ and $t < 0$, as illustrated in Figure 9.9(a). For such a signal, the ROC can be examined by choosing an arbitrary time T_0 and dividing $x(t)$ into the sum of a right-sided signal $x_R(t)$ and a left-sided signal $x_L(t)$, as indicated in Figures 9.9(b) and 9.9(c). The Laplace transform of $x(t)$ converges for values of s for which the transforms of *both* $x_R(t)$ and $x_L(t)$ converge. From Property 4, the ROC of $\mathcal{L}\{x_R(t)\}$ consists of a half-plane $\mathcal{R}e\{s\} > \sigma_R$ for some value σ_R, and from Property 5, the ROC of $\mathcal{L}\{x_L(t)\}$ consists of a half-plane $\mathcal{R}e\{s\} < \sigma_L$ for some value σ_L. The ROC of $\mathcal{L}\{x(t)\}$ is then the overlap of these two half-planes, as indicated in Figure 9.10. This assumes, of course, that $\sigma_R < \sigma_L$, so that there is some overlap. If this is not the case, then even if the Laplace transforms of $x_R(t)$ and $x_L(t)$ individually exist, the Laplace transform of $x(t)$ does not.

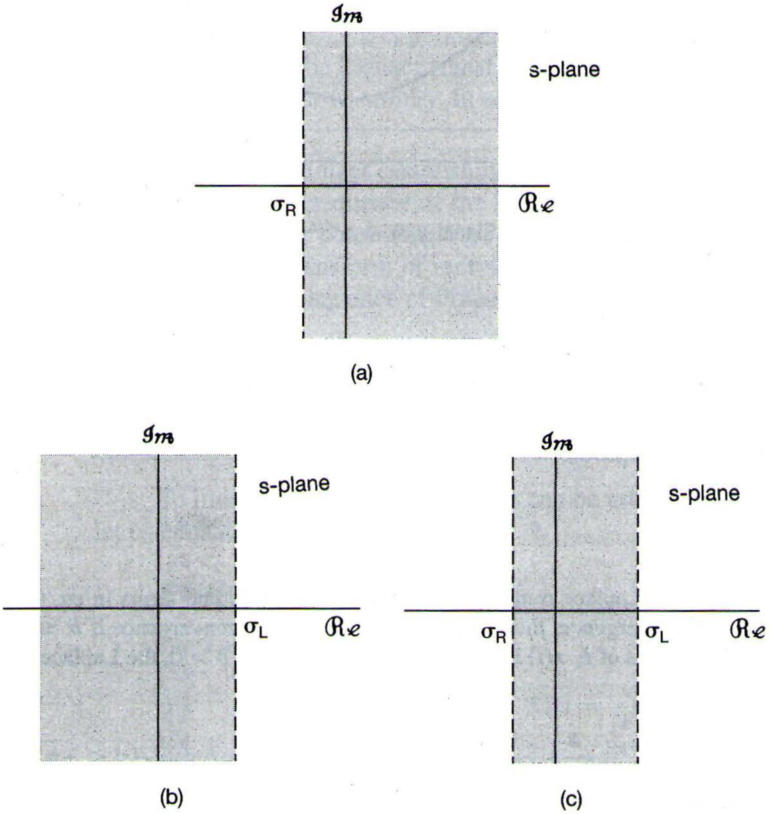

x(t)
T0
t
xR(t)
T0
t
xL(t)
T0
t

(a)

(b)

(c)

Figure 9.9 Two-sided signal divided into the sum of a right-sided and left-sided signal: (a) two-sided signal $x(t)$; (b) the right-sided signal equal to $x(t)$ for $t > T_0$ and equal to 0 for $t < T_0$; (c) the left-sided signal equal to $x(t)$ for $t < T_0$ and equal to 0 for $t > T_0$.


Im
s-plane
σR
Re
Im
s-plane
σL
Re
Im
s-plane
σR
σL
Re

(a)

(b)

(c)

Figure 9.10 (a) ROC for $x_R(t)$ in Figure 9.9; (b) ROC for $x_L(t)$ in Figure 9.9; (c) the ROC for $x(t) = x_R(t) + x_L(t)$, assuming that the ROCs in (a) and (b) overlap.

Example 9.7

Let

$$x(t) = e^{-b|t|}, \tag{9.47}$$

as illustrated in Figure 9.11 for both $b > 0$ and $b < 0$. Since this is a two-sided signal, let us divide it into the sum of a right-sided and left-sided signal; that is,

$$x(t) = e^{-bt}u(t) + e^{+bt}u(-t). \tag{9.48}$$

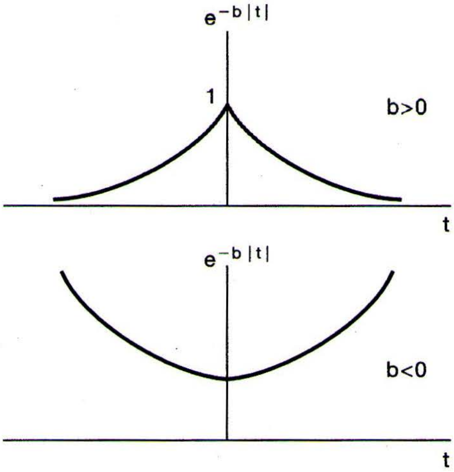


Figure 9.11 Signal $x(t) = e^{-b|t|}$ for both $b > 0$ and $b < 0$.

From Example 9.1,

$$e^{-bt}u(t) \overset{\mathcal{L}}{\longleftrightarrow} \frac{1}{s+b}, \quad \mathcal{R}e\{s\} > -b, \tag{9.49}$$

and from Example 9.2,

$$e^{+bt}u(-t) \overset{\mathcal{L}}{\longleftrightarrow} \frac{-1}{s-b}, \quad \mathcal{R}e\{s\} < +b. \tag{9.50}$$

Although the Laplace transforms of each of the individual terms in eq. (9.48) have a region of convergence, there is no *common* region of convergence if $b \le 0$, and thus, for those values of b, $x(t)$ has no Laplace transform. If $b > 0$, the Laplace transform of $x(t)$ is

$$e^{-b|t|} \overset{\mathcal{L}}{\longleftrightarrow} \frac{1}{s+b} - \frac{1}{s-b} = \frac{-2b}{s^2-b^2}, \quad -b < \mathcal{R}e\{s\} < +b. \tag{9.51}$$

The corresponding pole-zero plot is shown in Figure 9.12, with the shading indicating the ROC.

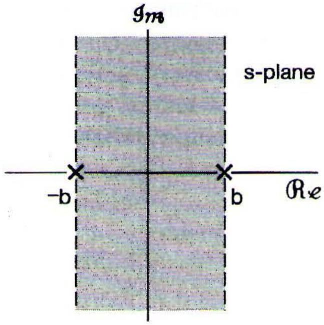


Figure 9.12 Pole-zero plot and ROC for Example 9.7.

A signal either does not have a Laplace transform or falls into one of the four categories covered by Properties 3 through 6. Thus, for any signal with a Laplace transform, the ROC *must* be the entire s-plane (for finite-length signals), a left-half plane (for left-sided signals), a right-half plane (for right-sided signals), or a single strip (for two-sided signals). In all the examples that we have considered, the ROC has the additional property that in each direction (i.e., $\mathcal{Re}\{s\}$ increasing and $\mathcal{Re}\{s\}$ decreasing) it is bounded by poles or extends to infinity. In fact, this is *always* true for rational Laplace transforms:

Property 7: If the Laplace transform $X(s)$ of $x(t)$ is rational, then its ROC is bounded by poles or extends to infinity. In addition, no poles of $X(s)$ are contained in the ROC.

A formal argument establishing this property is somewhat involved, but its validity is essentially a consequence of the facts that a signal with a rational Laplace transform consists of a linear combination of exponentials and, from Examples 9.1 and 9.2, that the ROC for the transform of individual terms in this linear combination must have the property. As a consequence of Property 7, together with Properties 4 and 5, we have

Property 8: If the Laplace transform $X(s)$ of $x(t)$ is rational, then if $x(t)$ is right sided, the ROC is the region in the s-plane to the right of the rightmost pole. If $x(t)$ is left sided, the ROC is the region in the s-plane to the left of the leftmost pole.

To illustrate how different ROCs can be associated with the same pole-zero pattern, let us consider the following example:

Example 9.8

Let

$$X(s) = \frac{1}{(s+1)(s+2)}, \tag{9.52}$$

with the associated pole-zero pattern in Figure 9.13(a). As indicated in Figures 9.13(b)–(d), there are three possible ROCs that can be associated with this algebraic expression, corresponding to three distinct signals. The signal associated with the pole-zero pattern in Figure 9.13(b) is right sided. Since the ROC includes the $j\omega$-axis, the Fourier

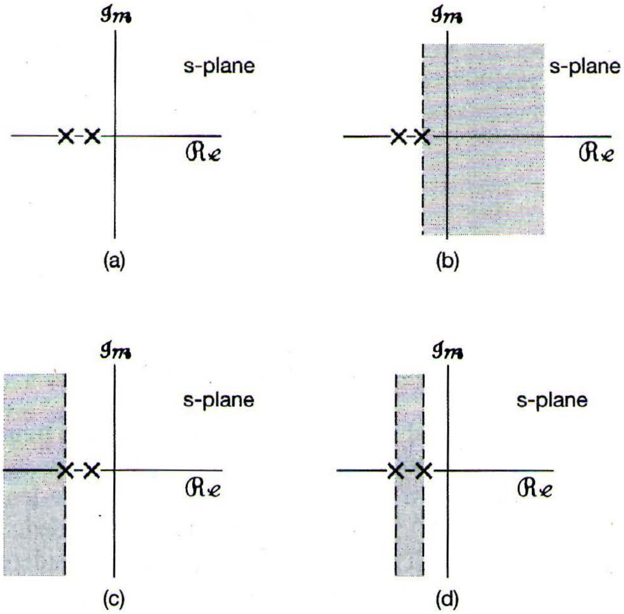


Figure 9.13 (a) Pole-zero pattern for Example 9.8; (b) ROC corresponding to a right-sided sequence; (c) ROC corresponding to a left-sided sequence; (d) ROC corresponding to a two-sided sequence.

transform of this signal converges. Figure 9.13(c) corresponds to a left-sided signal and Figure 9.13(d) to a two-sided signal. Neither of these two signals have Fourier transforms, since their ROCs do not include the $j\omega$-axis.

9.3 THE INVERSE LAPLACE TRANSFORM

In Section 9.1 we discussed the interpretation of the Laplace transform of a signal as the Fourier transform of an exponentially weighted version of the signal; that is, with s expressed as $s = \sigma + j\omega$, the Laplace transform of a signal $x(t)$ is

$$X(\sigma + j\omega) = \mathfrak{F}\{x(t)e^{-\sigma t}\} = \int_{-\infty}^{+\infty} x(t)e^{-\sigma t}e^{-j\omega t}\,dt \tag{9.53}$$

for values of $s = \sigma + j\omega$ in the ROC. We can invert this relationship using the inverse Fourier transform as given in eq. (4.9). We have

$$x(t)e^{-\sigma t} = \mathfrak{F}^{-1}\{X(\sigma + j\omega)\} = \frac{1}{2\pi}\int_{-\infty}^{+\infty} X(\sigma + j\omega)e^{j\omega t}\,d\omega, \tag{9.54}$$

or, multiplying both sides by $e^{\sigma t}$, we obtain

$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} X(\sigma + j\omega)e^{(\sigma + j\omega)t}\,d\omega. \tag{9.55}$$

That is, we can recover $x(t)$ from its Laplace transform evaluated for a set of values of $s = \sigma + j\omega$ in the ROC, with σ fixed and ω varying from $-\infty$ to $+\infty$. We can highlight this and gain additional insight into recovering $x(t)$ from $X(s)$ by changing the variable of integration in eq. (9.55) from ω to s and using the fact that σ is constant, so that $ds = j\,d\omega$. The result is the basic inverse Laplace transform equation:

$$\boxed{x(t) = \frac{1}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty} X(s)e^{st}\,ds.} \tag{9.56}$$

This equation states that $x(t)$ can be represented as a weighted integral of complex exponentials. The contour of integration in eq. (9.56) is the straight line in the s-plane corresponding to all points s satisfying $\mathcal{R}e\{s\} = \sigma$. This line is parallel to the $j\omega$-axis. Furthermore, we can choose any such line in the ROC—i.e., we can choose any value of σ such that $X(\sigma + j\omega)$ converges. The formal evaluation of the integral for a general $X(s)$ requires the use of contour integration in the complex plane, a topic that we will not consider here. However, for the class of rational transforms, the inverse Laplace transform can be determined without directly evaluating eq. (9.56) by using the technique of partial-fraction expansion in a manner similar to that used in Chapter 4 to determine the inverse Fourier transform. Basically, the procedure consists of expanding the rational algebraic expression into a linear combination of lower order terms.

For example, assuming no multiple-order poles, and assuming that the order of the denominator polynomial is greater than the order of the numerator polynomial, we can expand $X(s)$ in the form

$$X(s) = \sum_{i=1}^{m} \frac{A_i}{s + a_i}. \tag{9.57}$$

From the ROC of $X(s)$, the ROC of each of the individual terms in eq. (9.57) can be inferred, and then, from Examples 9.1 and 9.2, the inverse Laplace transform of each of these terms can be determined. There are two possible choices for the inverse transform of each term $A_i/(s + a_i)$ in the equation. If the ROC is to the right of the pole at $s = -a_i$, then the inverse transform of this term is $A_i e^{-a_i t}u(t)$, a right-sided signal. If the ROC is to the left of the pole at $s = -a_i$, then the inverse transform of the term is $-A_i e^{-a_i t}u(-t)$, a left-sided signal. Adding the inverse transforms of the individual terms in eq. (9.57) then yields the inverse transform of $X(s)$. The details of this procedure are best presented through a number of examples.

Example 9.9

Let

$$X(s) = \frac{1}{(s+1)(s+2)}, \qquad \mathcal{R}e\{s\} > -1. \tag{9.58}$$

To obtain the inverse Laplace transform, we first perform a partial-fraction expansion to obtain

$$X(s) = \frac{1}{(s+1)(s+2)} = \frac{A}{s+1} + \frac{B}{s+2}. \tag{9.59}$$

As discussed in the appendix, we can evaluate the coefficients A and B by multiplying both sides of eq. (9.59) by $(s + 1)(s + 2)$ and then equating coefficients of equal powers of s on both sides. Alternatively, we can use the relation

$$A = [(s + 1)X(s)]\big|_{s=-1} = 1, \tag{9.60}$$

$$B = [(s + 2)X(s)]\big|_{s=-2} = -1. \tag{9.61}$$

Thus, the partial-fraction expansion for $X(s)$ is

$$X(s) = \frac{1}{s+1} - \frac{1}{s+2}. \tag{9.62}$$

From Examples 9.1 and 9.2, we know that there are two possible inverse transforms for a transform of the form $1/(s + a)$, depending on whether the ROC is to the left or the right of the pole. Consequently, we need to determine which ROC to associate with each of the individual first-order terms in eq. (9.62). This is done by reference to the properties of the ROC developed in Section 9.2. Since the ROC for $X(s)$ is $\mathcal{R}e\{s\} > -1$, the ROC for the individual terms in the partial-fraction expansion of eq. (9.62) includes $\mathcal{R}e\{s\} > -1$. The ROC for each term can then be extended to the left or right (or both) to be bounded by a pole or infinity. This is illustrated in Figure 9.14. Figure 9.14(a) shows the pole-zero plot and ROC for $X(s)$, as specified in eq. (9.58). Figure 9.14(b) and 9.14(c) represent the individual terms in the partial-fraction expansion in eq. (9.62). The ROC for the sum is indicated with lighter shading. For the term represented by Figure 9.14(c), the ROC for the sum can be extended to the left as shown, so that it is bounded by a pole.

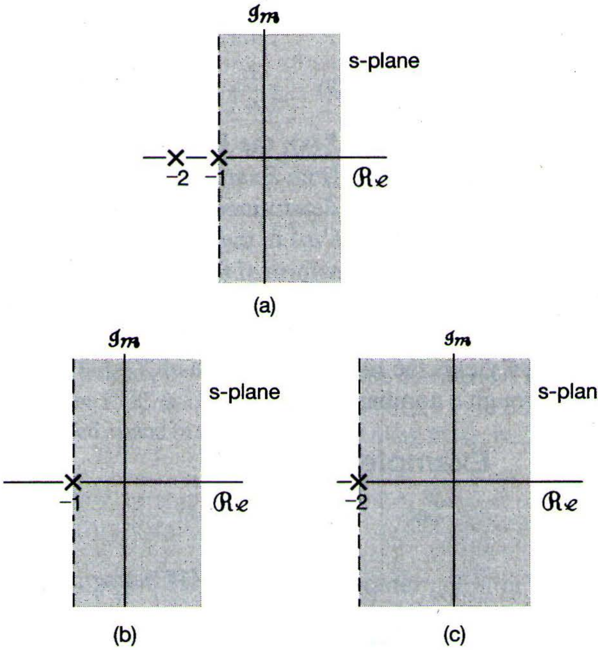


Figure 9.14 Construction of the ROCs for the individual terms in the partial-fraction expansion of $X(s)$ in Example 9.8: (a) pole-zero plot and ROC for $X(s)$; (b) pole at $s = -1$ and its ROC; (c) pole at $s = -2$ and its ROC.

Since the ROC is to the right of both poles, the same is true for each of the individual terms, as can be seen in Figures 9.14(b) and (c). Consequently, from Property 8 in the preceding section, we know that each of these terms corresponds to a right-sided signal. The inverse transform of the individual terms in eq. (9.62) can then be obtained by reference to Example 9.1:

$$e^{-t}u(t) \overset{\mathcal{L}}{\longleftrightarrow} \frac{1}{s+1}, \qquad \mathcal{R}e\{s\} > -1, \tag{9.63}$$

$$e^{-2t}u(t) \overset{\mathcal{L}}{\longleftrightarrow} \frac{1}{s+2}, \qquad \mathcal{R}e\{s\} > -2. \tag{9.64}$$

We thus obtain

$$[e^{-t} - e^{-2t}]u(t) \overset{\mathcal{L}}{\longleftrightarrow} \frac{1}{(s+1)(s+2)}, \qquad \mathcal{R}e\{s\} > -1. \tag{9.65}$$

Example 9.10

Let us now suppose that the algebraic expression for $X(s)$ is again given by eq. (9.58), but that the ROC is now the left-half plane $\mathcal{R}e\{s\} < -2$. The partial-fraction expansion for $X(s)$ relates only to the algebraic expression, so eq. (9.62) is still valid. With this new ROC, however, the ROC is to the *left* of both poles and thus, the same must be true for each of the two terms in the equation. That is, the ROC for the term corresponding to the pole at $s = -1$ is $\mathcal{R}e\{s\} < -1$, while the ROC for the term with pole at $s = -2$ is $\mathcal{R}e\{s\} < -2$. Then, from Example 9.2,

$$-e^{-t}u(-t) \overset{\mathcal{L}}{\longleftrightarrow} \frac{1}{s+1}, \qquad \mathcal{R}e\{s\} < -1, \tag{9.66}$$

$$-e^{-2t}u(-t) \overset{\mathcal{L}}{\longleftrightarrow} \frac{1}{s+2}, \qquad \mathcal{R}e\{s\} < -2, \tag{9.67}$$

so that

$$x(t) = [-e^{-t} + e^{-2t}]u(-t) \overset{\mathcal{L}}{\longleftrightarrow} \frac{1}{(s+1)(s+2)}, \qquad \mathcal{R}e\{s\} < -2. \tag{9.68}$$

Example 9.11

Finally, suppose that the ROC of $X(s)$ in eq. (9.58) is $-2 < \mathcal{R}e\{s\} < -1$. In this case, the ROC is to the left of the pole at $s = -1$ so that this term corresponds to the left-sided signal in eq. (9.66), while the ROC is to the right of the pole at $s = -2$ so that this term corresponds to the right-sided signal in eq. (9.64). Combining these, we find that

$$x(t) = -e^{-t}u(-t) - e^{-2t}u(t) \overset{\mathcal{L}}{\longleftrightarrow} \frac{1}{(s+1)(s+2)}, \qquad -2 < \mathcal{R}e\{s\} < -1. \tag{9.69}$$

As discussed in the appendix, when $X(s)$ has multiple-order poles, or when the denominator is not of higher degree than the numerator, the partial-fraction expansion of $X(s)$ will include other terms in addition to the first-order terms considered in Examples 9.9–9.11. In Section 9.5, after discussing properties of the Laplace transform, we develop some other Laplace transform pairs that, in conjunction with the properties, allow us to extend the inverse transform method outlined in Example 9.9 to arbitrary rational transforms.

9.4 GEOMETRIC EVALUATION OF THE FOURIER TRANSFORM FROM THE POLE-ZERO PLOT

As we saw in Section 9.1, the Fourier transform of a signal is the Laplace transform evaluated on the $j\omega$-axis. In this section we discuss a procedure for geometrically evaluating the Fourier transform and, more generally, the Laplace transform at any set of values from the pole-zero pattern associated with a rational Laplace transform. To develop the procedure, let us first consider a Laplace transform with a single zero [i.e., $X(s) = s - a$], which we evaluate at a specific value of s, say, $s = s_1$. The algebraic expression $s_1 - a$ is the sum of two complex numbers, s_1 and $-a$, each of which can be represented as a vector in the complex plane, as illustrated in Figure 9.15. The vector representing the complex number $s_1 - a$ is then the vector sum of $\mathbf{s}_1$ and $-\mathbf{a}$, which we see in the figure to be a vector from the zero at $s = a$ to the point s_1. The value of $X(s_1)$ then has a magnitude that is the length of this vector and an angle that is the angle of the vector relative to the real axis. If $X(s)$ instead has a single pole at $s = a$ [i.e., $X(s) = 1/(s - a)$], then the denominator would be represented by the same vector sum of $\mathbf{s}_1$ and $-\mathbf{a}$, and the value of $X(s_1)$ would have a magnitude that is the *reciprocal* of the length of the vector from the pole to $s = s_1$ and an angle that is the *negative* of the angle of the vector with the real axis.

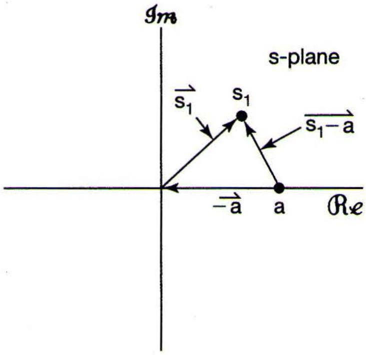


Figure 9.15 Complex plane representation of the vectors $\mathbf{s}_1$, $\mathbf{a}$, and $\mathbf{s}_1 - \mathbf{a}$ representing the complex numbers s_1, a, and $s_1 - a$, respectively.

A more general rational Laplace transform consists of a product of pole and zero terms of the form discussed in the preceding paragraph; that is, it can be factored into the form

$$X(s) = M \frac{\prod_{i=l}^{R}(s - \beta_i)}{\prod_{j=l}^{P}(s - \alpha_j)}. \tag{9.70}$$

To evaluate $X(s)$ at $s = s_1$, each term in the product is represented by a vector from the zero or pole to the point s_1. The magnitude of $X(s_1)$ is then the magnitude of the scale factor M, times the product of the lengths of the zero vectors (i.e., the vectors from the zeros to s_1) divided by the product of the lengths of the pole vectors (i.e., the vectors from the poles to s_1). The angle of the complex number $X(s_1)$ is the sum of the angles of the zero vectors minus the sum of the angles of the pole vectors. If the scale factor M in eq. (9.70) is negative, an additional angle of π would be included. If $X(s)$ has a multiple pole or zero (or both), corresponding to some of the α_j's being equal to each other or some of the β_i's being equal to each other (or both), the lengths and angles of the vectors from each of these poles or zeros must be included a number of times equal to the order of the pole or zero.

Example 9.12

Let

$$X(s) = \frac{1}{s+\frac{1}{2}}, \qquad \mathcal{R}e\{s\} > -\frac{1}{2}. \tag{9.71}$$

The Fourier transform is $X(s)|_{s=j\omega}$. For this example, then, the Fourier transform is

$$X(j\omega) = \frac{1}{j\omega + 1/2}. \tag{9.72}$$

The pole-zero plot for $X(s)$ is shown in Figure 9.16. To determine the Fourier transform graphically, we construct the pole vector as indicated. The magnitude of the Fourier transform at frequency ω is the reciprocal of the length of the vector from the pole to the point $j\omega$ on the imaginary axis. The phase of the Fourier transform is the negative of the angle of the vector. Geometrically, from Figure 9.16, we can write

$$|X(j\omega)|^2 = \frac{1}{\omega^2 + (1/2)^2} \tag{9.73}$$

and

$$\sphericalangle X(j\omega) = -\tan^{-1} 2\omega. \tag{9.74}$$

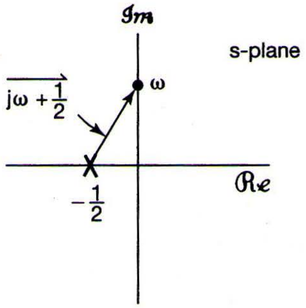


Figure 9.16 Pole-zero plot for Example 9.12. $|X(j\omega)|$ is the reciprocal of the length of the vector shown, and $\sphericalangle X(j\omega)$ is the negative of the angle of the vector.

Often, part of the value of the geometric determination of the Fourier transform lies in its usefulness in obtaining an approximate view of the overall characteristics of the transform. For example, in Figure 9.16, it is readily evident that the length of the pole vector monotonically increases with increasing ω, and thus, the magnitude of the Fourier transform will monotonically *decrease* with increasing ω. The ability to draw general conclusions about the behavior of the Fourier transform from the pole-zero plot is further illustrated by a consideration of general first- and second-order systems.

9.4.1 First-Order Systems

As a generalization of Example 9.12, let us consider the class of first-order systems that was discussed in some detail in Section 6.5.1. The impulse response for such a system is

$$h(t) = \frac{1}{\tau}e^{-t/\tau}u(t), \tag{9.75}$$

and its Laplace transform is

$$H(s) = \frac{1}{s\tau + 1}, \qquad \mathcal{R}e\{s\} > -\frac{1}{\tau}. \tag{9.76}$$

The pole-zero plot is shown in Figure 9.17. Note from the figure that the length of the pole vector is minimal for $\omega = 0$ and increases monotonically as ω increases. Also, the angle of the pole increases monotonically from 0 to $\pi/2$ as ω increases from 0 to ∞.

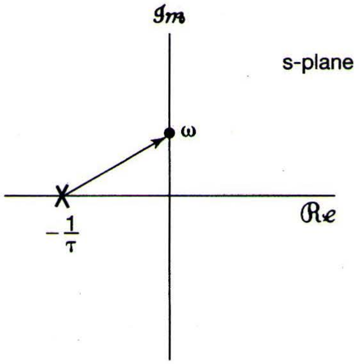


Figure 9.17 Pole-zero plot for first-order system of eq. (9.76).

From the behavior of the pole vector as ω varies, it is clear that the magnitude of the frequency response $H(j\omega)$ monotonically decreases as ω increases, while $\sphericalangle H(j\omega)$ monotonically decreases from 0 to $-\pi/2$, as shown in the Bode plots for this system in Figure 9.18. Note also that when $\omega = 1/\tau$, the real and imaginary parts of the pole vector are equal, yielding a value of the magnitude of the frequency response that is reduced by a factor of $\sqrt{2}$, or approximately 3 dB, from its maximum at $\omega = 0$ and a value of $\pi/4$ for the angle of the frequency response. This is consistent with our examination of first-order systems in Section 6.5.1, where we noted that $\omega = 1/\tau$ is often referred to as the 3-dB point or the break frequency—i.e., the frequency at which the straight-line approximation of the Bode plot of $|H(j\omega)|$ has a break in its slope. As we also saw in Section 6.5.1, the time constant τ controls the speed of response of first-order systems, and we now see that

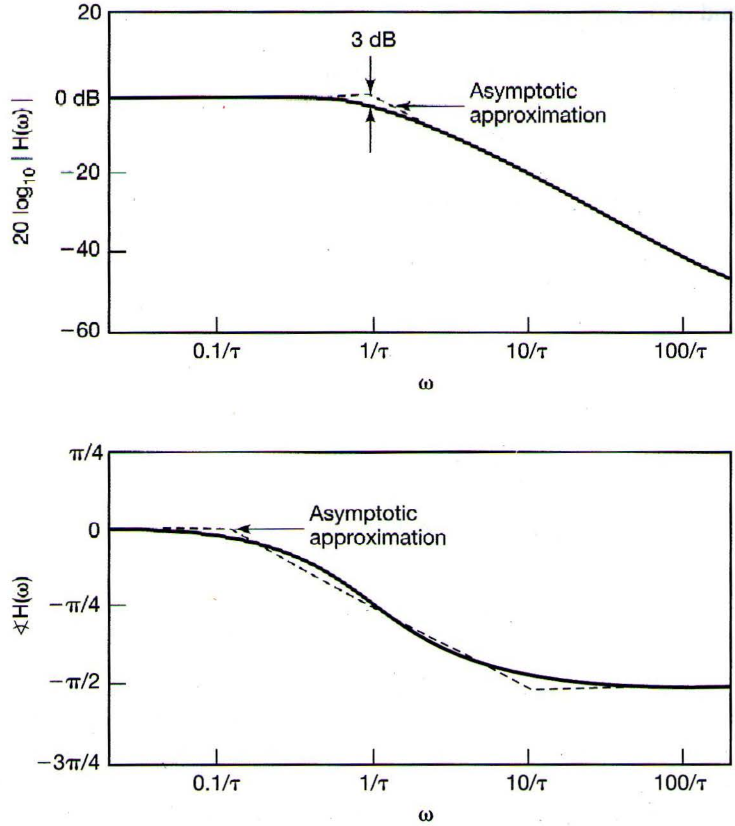


Figure 9.18 Frequency response for a first-order system.

the pole of such a system at $s = -1/\tau$ is on the negative real axis, at a distance to the origin that is the reciprocal of the time constant.

From our graphical interpretation, we can also see how changing the time constant or, equivalently, the position of the pole of $H(s)$ changes the characteristics of a first-order system. In particular, as the pole moves farther into the left-half plane, the break frequency and, hence, the effective cutoff frequency of the system increases. Also, from eq. (9.75) and from Figure 6.19, we see that this same movement of the pole to the left corresponds to a decrease in the time constant τ, resulting in a faster decay of the impulse response and a correspondingly faster rise time in the step response. This relationship between the real part of the pole locations and the speed of the system response holds more generally; that is, poles farther away from the $j\omega$-axis are associated with faster response terms in the impulse response.

9.4.2 Second-Order Systems

Let us next consider the class of second-order systems, which was discussed in some detail in Section 6.5.2. The impulse response and frequency response for the system, originally given in eqs. (6.37) and (6.33), respectively, are

$$h(t) = M[e^{c_1 t} - e^{c_2 t}]u(t), \tag{9.77}$$

where

$$\begin{aligned} c_1 &= -\zeta\omega_n + \omega_n\sqrt{\zeta^2 - 1}, \\ c_2 &= -\zeta\omega_n - \omega_n\sqrt{\zeta^2 - 1}, \\ M &= \frac{\omega_n}{2\sqrt{\zeta^2 - 1}}, \end{aligned}$$

and

$$H(j\omega) = \frac{\omega_n^2}{(j\omega)^2 + 2\zeta\omega_n(j\omega) + \omega_n^2}. \tag{9.78}$$

The Laplace transform of the impulse response is

$$H(s) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} = \frac{\omega_n^2}{(s - c_1)(s - c_2)}. \tag{9.79}$$

For $\zeta > 1$, c_1 and c_2 are real and thus both poles lie on the real axis, as indicated in Figure 9.19(a). The case of $\zeta > 1$ is essentially a product of two first-order terms, as in Section 9.4.1. Consequently, in this case $|H(j\omega)|$ decreases monotonically as $|\omega|$ increases, while $\sphericalangle H(j\omega)$ varies from 0 at $\omega = 0$ to $-\pi$ as $\omega \rightarrow \infty$. This can be verified from Figure 9.19(a) by observing that the length of the vector from each of the two poles to the point $s = j\omega$ increases monotonically as ω increases from 0, and the angle of each of these vectors increases from 0 to $\pi/2$ as ω increases from 0 to ∞. Note also that as ζ increases, one pole moves closer to the $j\omega$-axis, indicative of a term in the impulse response that decays more slowly, and the other pole moves farther into the left-half plane, indicative of a term in the impulse response that decays more rapidly. Thus, for large values of ζ, it is the pole close to the $j\omega$-axis that dominates the system response for large time. Similarly, from a consideration of the pole vectors for $\zeta \gg 1$, as indicated in Figure 9.19(b), for low frequencies the length and angle of the vector for the pole close to the $j\omega$-axis are much more sensitive to changes in ω than the length and angle of the vector for the pole far from the $j\omega$-axis. Hence, we see that for low frequencies, the characteristics of the frequency response are influenced principally by the pole close to the $j\omega$-axis.

For $0 < \zeta < 1$, c_1 and c_2 are complex, so that the pole-zero plot is that shown in Figure 9.19(c). Correspondingly, the impulse response and step response have oscillatory parts. We note that the two poles occur in complex conjugate locations. In fact, as we discuss in Section 9.5.5, the complex poles (and zeros) for a real-valued signal always occur in complex conjugate pairs. From the figure—particularly when ζ is small, so that the poles are close to the $j\omega$-axis—as ω approaches $\omega_n\sqrt{1 - \zeta^2}$, the behavior of the frequency response is dominated by the pole vector in the second quadrant, and in

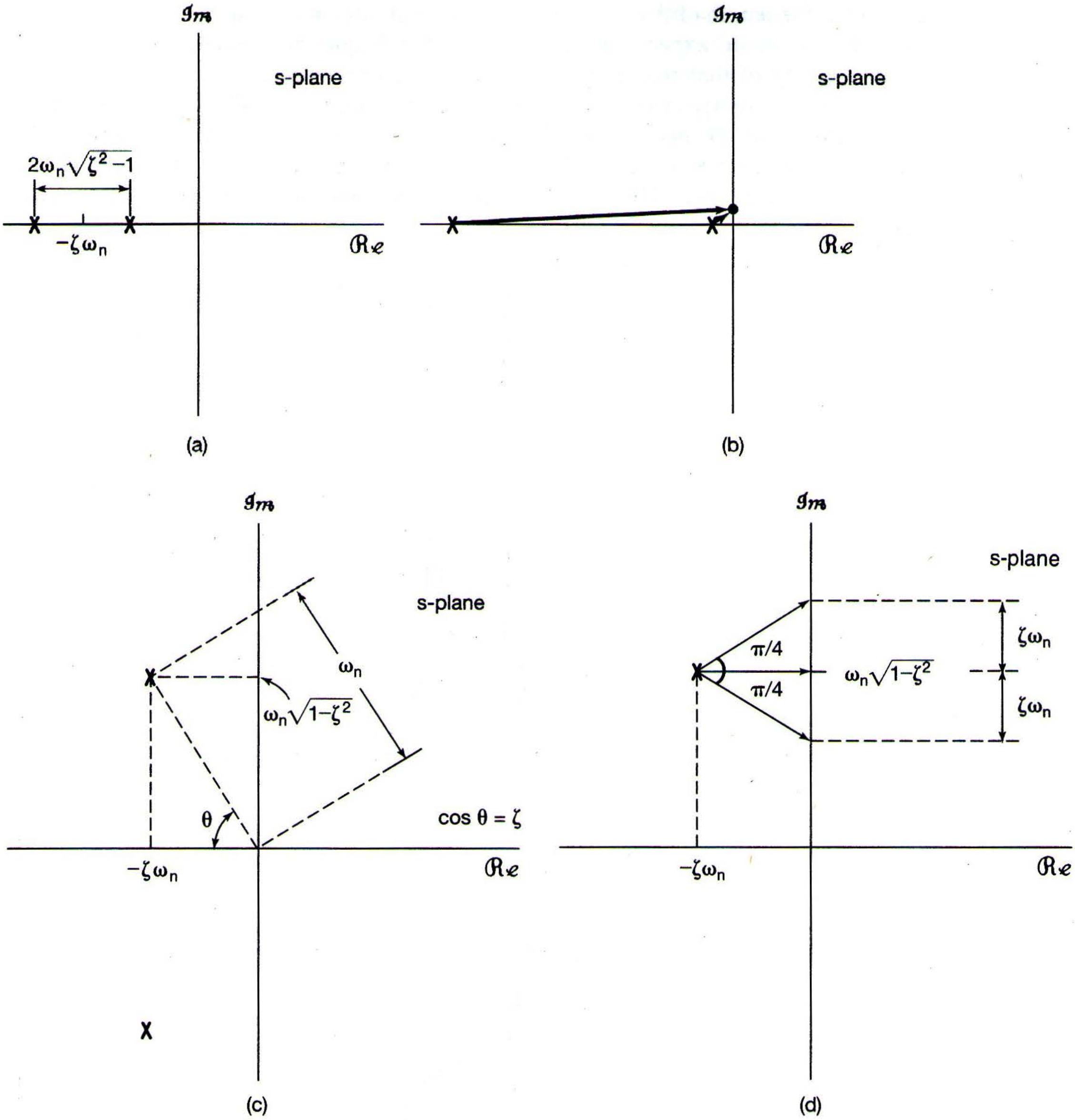

$\mathcal{Im}$
s-plane
$2\omega_n\sqrt{\zeta^2-1}$
$-\zeta\omega_n$
$\mathcal{Re}$
(a)
$\mathcal{Im}$
s-plane
$\mathcal{Re}$
(b)
$\mathcal{Im}$
s-plane
ω_n
$\omega_n\sqrt{1-\zeta^2}$
θ
$\cos\theta = \zeta$
$-\zeta\omega_n$
$\mathcal{Re}$
(c)
$\mathcal{Im}$
s-plane
$\pi/4$
$\pi/4$
$\omega_n\sqrt{1-\zeta^2}$
$\zeta\omega_n$
$\zeta\omega_n$
$-\zeta\omega_n$
$\mathcal{Re}$
(d)

Figure 9.19 (a) Pole-zero plot for a second-order system with $\zeta > 1$; (b) pole vectors for $\zeta \gg 1$; (c) pole-zero plot for a second-order system with $0 < \zeta < 1$; (d) pole vectors for $0 < \zeta < 1$ and for $\omega = \omega_n\sqrt{1-\zeta^2}$ and $\omega = \omega_n\sqrt{1-\zeta^2} \pm \zeta\omega_n$.

particular, the length of that pole vector has a minimum at $\omega = \omega_n\sqrt{1-\zeta^2}$. Thus, qualitatively, we would expect the magnitude of the frequency response to exhibit a peak in the vicinity of that frequency. Because of the presence of the other pole, the peak will occur not exactly at $\omega = \omega_n\sqrt{1-\zeta^2}$, but at a frequency slightly less than this. A careful sketch of the magnitude of the frequency response is shown in Figure 9.20(a) for $\omega_n = 1$ and several values of ζ where the expected behavior in the vicinity of the poles is clearly evident. This is consistent with our analysis of second-order systems in Section 6.5.2.

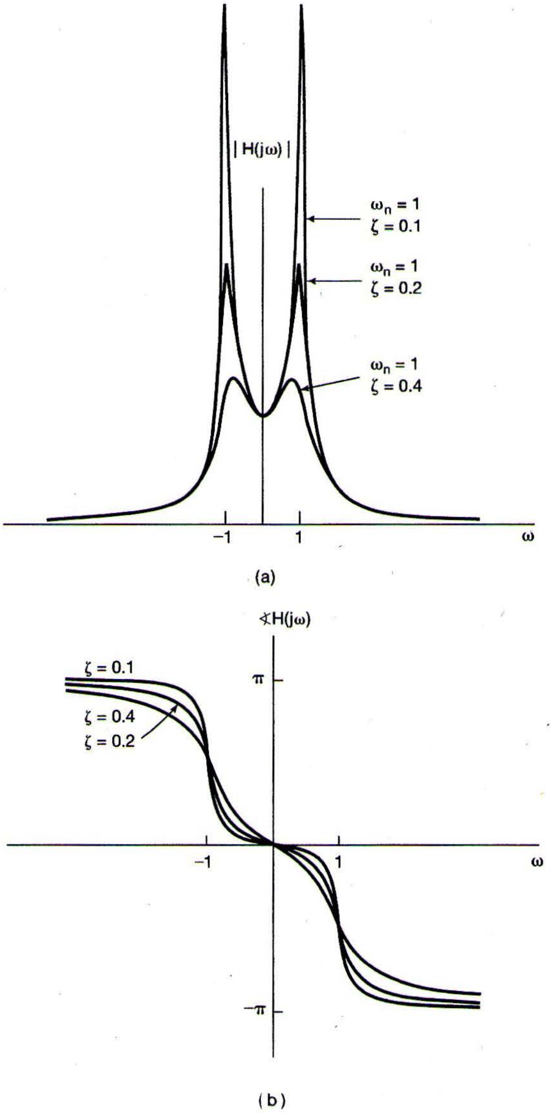


Figure 9.20 (a) Magnitude and (b) phase of the frequency response for a second-order system with $0 < \zeta < 1$.

Thus, for $0 < \zeta < 1$, the second-order system is a nonideal bandpass filter, with the parameter ζ controlling the sharpness and width of the peak in the frequency response. In particular, from the geometry in Figure 9.19(d), we see that the length of the pole vector from the second-quadrant pole increases by a factor of $\sqrt{2}$ from its minimum at $\omega = \omega_n\sqrt{1-\zeta^2}$ when ω increases or decreases from this value by $\zeta\omega_n$. Consequently, for small ζ, and neglecting the effect of the distant third-quadrant pole, $|H(j\omega)|$ is within a factor of $\sqrt{2}$ of its peak value over the frequency range

$$\omega_n\sqrt{1-\zeta^2} - \zeta\omega_n < \omega < \omega_n\sqrt{1-\zeta^2} + \zeta\omega_n.$$

If we define the relative bandwidth B as the length of this frequency interval divided by the undamped natural frequency ω_n, we see that

$$B = 2\zeta.$$

Thus, the closer ζ is to zero, the sharper and narrower the peak in the frequency response is. Note also that B is the reciprocal of the quality measure Q for second-order systems defined in Section 6.5.2. Thus, as the quality increases, the relative bandwidth decreases and the filter becomes increasingly frequency selective.

An analogous picture can be developed for $\sphericalangle H(\omega)$, which is plotted in Figure 9.20(b) for $\omega_n = 1$ and several values of ζ. As can be seen from Figure 9.19(d), the angle of the second-quadrant pole vector changes from $-\pi/4$ to 0 to $\pi/4$ as ω changes from $\omega_n\sqrt{1-\zeta^2} - \zeta\omega_n$ to $\omega_n\sqrt{1-\zeta^2}$ to $\omega_n\sqrt{1-\zeta^2} + \zeta\omega_n$. For small values of ζ, the angle for the third-quadrant pole changes very little over this frequency interval, resulting in a rapid change in $\sphericalangle H(j\omega)$ of approximately $\pi/2$ over the interval, as captured in the figure.

Varying ω_n with ζ fixed only changes the frequency scale in the preceding discussion—i.e., $|H(\omega)|$ and $\sphericalangle H(\omega)$ depend only on ω/ω_n. From Figure 9.19(c), we also can readily determine how the poles and system characteristics change as we vary ζ, keeping ω_n constant. Since $\cos\theta = \zeta$, the poles move along a semicircle with fixed radius ω_n. For $\zeta = 0$, the two poles are on the imaginary axis. Correspondingly, in the time domain, the impulse response is sinusoidal with no damping. As ζ increases from 0 to 1, the two poles remain complex and move into the left-half plane, and the vectors from the origin to the poles maintain a constant overall magnitude ω_n. As the real part of the poles becomes more negative, the associated time response will decay more quickly as $t \to \infty$. Also, as we have seen, as ζ increases from 0 toward 1, the relative bandwidth of the frequency response increases, and the frequency response becomes less sharp and less frequency selective.

9.4.3 All-Pass Systems

As a final illustration of the geometric evaluation of the frequency response, let us consider a system for which the Laplace transform of the impulse response has the pole-zero plot shown in Figure 9.21(a). From this figure, it is evident that for any point along the $j\omega$-axis, the pole and zero vectors have equal length, and consequently, the magnitude of the frequency response is constant and independent of frequency. Such a system is com-

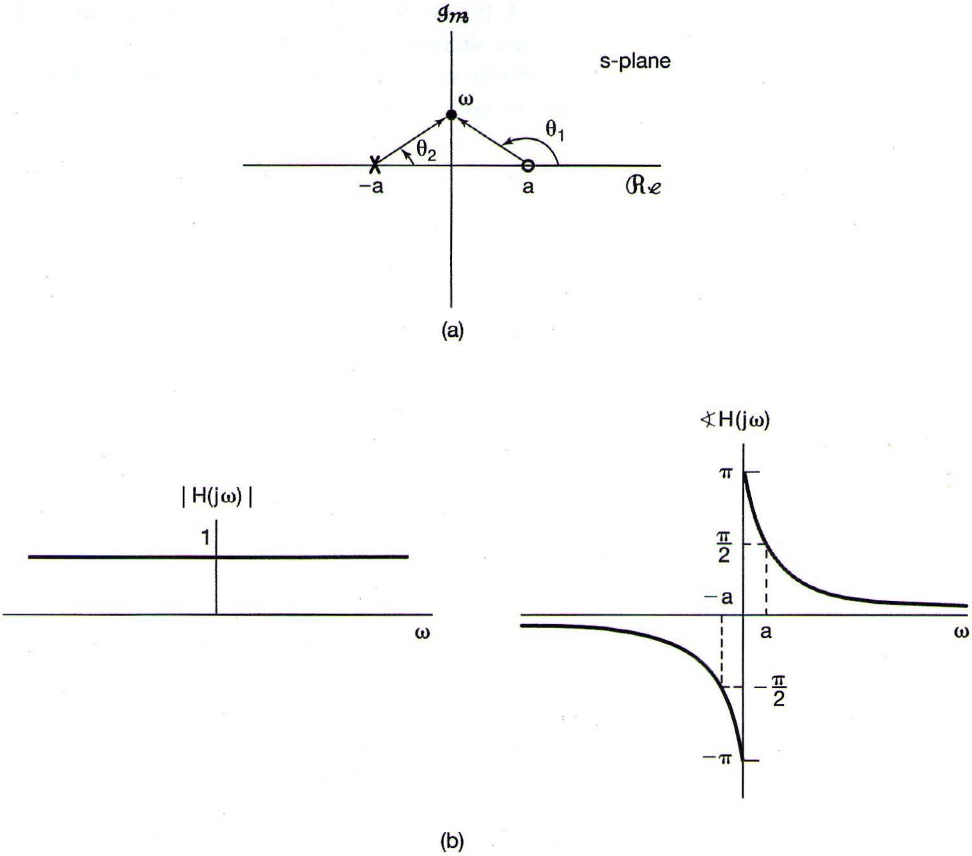


Figure 9.21 (a) Pole-zero plot for an all-pass system; (b) magnitude and phase of an all-pass frequency response.

monly referred to as an *all-pass system,* since it passes all frequencies with equal gain (or attenuation). The phase of the frequency response is $\theta_1 - \theta_2$, or, since $\theta_1 = \pi - \theta_2$,

$$\sphericalangle H(j\omega) = \pi - 2\theta_2. \tag{9.80}$$

From Figure 9.21(a), $\theta_2 = \tan^{-1}(\omega/a)$, and thus,

$$\sphericalangle H(j\omega) = \pi - 2\tan^{-1}\left(\frac{\omega}{a}\right). \tag{9.81}$$

The magnitude and phase of $H(j\omega)$ are illustrated in Figure 9.21(b).

9.5 PROPERTIES OF THE LAPLACE TRANSFORM

In exploiting the Fourier transform, we relied heavily on the set of properties developed in Section 4.3. In the current section, we consider the corresponding set of properties for the Laplace transform. The derivations of many of these results are analogous to those of the corresponding properties for the Fourier transform. Consequently, we will not present the derivations in detail, some of which are left as exercises at the end of the chapter. (See Problems 9.52–9.54.)

9.5.1 Linearity of the Laplace Transform

If

$$x_1(t) \overset{\mathcal{L}}{\longleftrightarrow} X_1(s) \quad \text{with a region of convergence that will be denoted as } R_1$$

and

$$x_2(t) \overset{\mathcal{L}}{\longleftrightarrow} X_2(s) \quad \text{with a region of convergence that will be denoted as } R_2,$$

then

$$\boxed{ax_1(t) + bx_2(t) \overset{\mathcal{L}}{\longleftrightarrow} aX_1(s) + bX_2(s), \text{ with ROC containing } R_1 \cap R_2.} \tag{9.82}$$

As indicated, the region of convergence of $X(s)$ is at least the intersection of R_1 and R_2, which could be empty, in which case $X(s)$ has no region of convergence—i.e., $x(t)$ has no Laplace transform. For example, for $x(t)$ as in eq. (9.47) of Example 9.7, with $b > 0$ the ROC for $X(s)$ is the intersection of the ROCs for the two terms in the sum. If $b < 0$, there are no common points in R_1 and R_2; that is, the intersection is empty, and thus, $x(t)$ has no Laplace transform. The ROC can also be larger than the intersection. As a simple example, for $x_1(t) = x_2(t)$ and $a = -b$ in eq. (9.82), $x(t) = 0$, and thus, $X(s) = 0$. The ROC of $X(s)$ is then the entire s-plane.

The ROC associated with a linear combination of terms can always be constructed by using the properties of the ROC developed in Section 9.2. Specifically, from the intersection of the ROCs for the individual terms (assuming that it is not empty), we can find a line or strip that is in the ROC of the linear combination. We then extend this to the right ($\mathcal{R}e\{s\}$ increasing) and to the left ($\mathcal{R}e\{s\}$ decreasing) to the nearest poles (which may be at infinity).

Example 9.13

In this example, we illustrate the fact that the ROC for the Laplace transform of a linear combination of signals can sometimes extend beyond the intersection of the ROCs for the individual terms. Consider

$$x(t) = x_1(t) - x_2(t), \tag{9.83}$$

where the Laplace transforms of $x_1(t)$ and $x_2(t)$ are, respectively,

$$X_1(s) = \frac{1}{s+1}, \qquad \Re e\{s\} > -1, \tag{9.84}$$

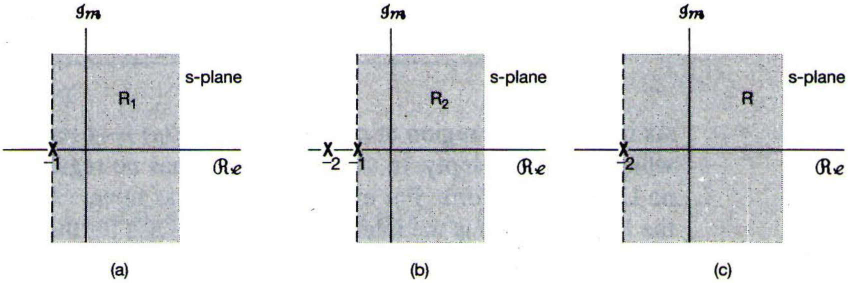
and

$$X_2(s) = \frac{1}{(s+1)(s+2)}, \qquad \Re e\{s\} > -1. \tag{9.85}$$

The pole-zero plot, including the ROCs for $X_1(s)$ and $X_2(s)$, is shown in Figures 9.22(a) and (b). From eq. (9.82),

$$X(s) = \frac{1}{s+1} - \frac{1}{(s+1)(s+2)} = \frac{s+1}{(s+1)(s+2)} = \frac{1}{s+2}. \tag{9.86}$$

Thus, in the linear combination of $x_1(t)$ and $x_2(t)$, the pole at $s = -1$ is canceled by a zero at $s = -1$. The pole-zero plot for $X(s) = X_1(s) - X_2(s)$ is shown in Figure 9.22(c). The intersection of the ROCs for $X_1(s)$ and $X_2(s)$ is $\Re e\{s\} > -1$. However, since the ROC is always bounded by a pole or infinity, for this example the ROC for $X(s)$ can be extended to the left to be bounded by the pole at $s = -2$, as a result of the pole-zero cancellation at $s = -1$.



(a) (b) (c)

Figure 9.22 Pole-zero plots and ROCs for Example 9.13: (a) $X_1(s)$; (b) $X_2(s)$; (c) $X_1(s) - X_2(s)$. The ROC for $X_1(s) - X_2(s)$ includes the intersection of R_1 and R_2, which can then be extended to be bounded by the pole at $s = -2$.

9.5.2 Time Shifting

If

$$x(t) \overset{\mathcal{L}}{\longleftrightarrow} X(s), \qquad \text{with ROC} = R,$$

then

$$\boxed{x(t-t_0) \overset{\mathcal{L}}{\longleftrightarrow} e^{-st_0}X(s), \quad \text{with ROC} = R.} \tag{9.87}$$

9.5.3 Shifting in the s-Domain

If

$$x(t) \overset{\mathcal{L}}{\longleftrightarrow} X(s), \qquad \text{with ROC} = R,$$

then

$$\boxed{e^{s_0 t}x(t) \overset{\mathcal{L}}{\longleftrightarrow} X(s - s_0), \quad \text{with ROC} = R + \mathcal{R}e\{s_0\}.} \tag{9.88}$$

That is, the ROC associated with $X(s - s_0)$ is that of $X(s)$, shifted by $\mathcal{R}e\{s_0\}$. Thus, for any value s that is in R, the value $s + \mathcal{R}e\{s_0\}$ will be in R_1. This is illustrated in Figure 9.23. Note that if $X(s)$ has a pole or zero at $s = a$, then $X(s-s_0)$ has a pole or zero at $s-s_0 = a$—i.e., $s = a + s_0$.

An important special case of eq. (9.88) is when $s_0 = j\omega_0$—i.e., when a signal $x(t)$ is used to modulate a periodic complex exponential $e^{j\omega_0 t}$. In this case, eq. (9.88) becomes

$$e^{j\omega_0 t}x(t) \overset{\mathcal{L}}{\longleftrightarrow} X(s - j\omega_0), \quad \text{with ROC} = R. \tag{9.89}$$

The right-hand side of eq. (9.89) can be interpreted as a shift in the s-plane parallel to the $j\omega$-axis. That is, if the Laplace transform of $x(t)$ has a pole or zero at $s = a$, then the Laplace transform of $e^{j\omega_0 t}x(t)$ has a pole or zero at $s = a + j\omega_0$.

9.5.4 Time Scaling

If

$$x(t) \overset{\mathcal{L}}{\longleftrightarrow} X(s), \quad \text{with ROC} = R,$$

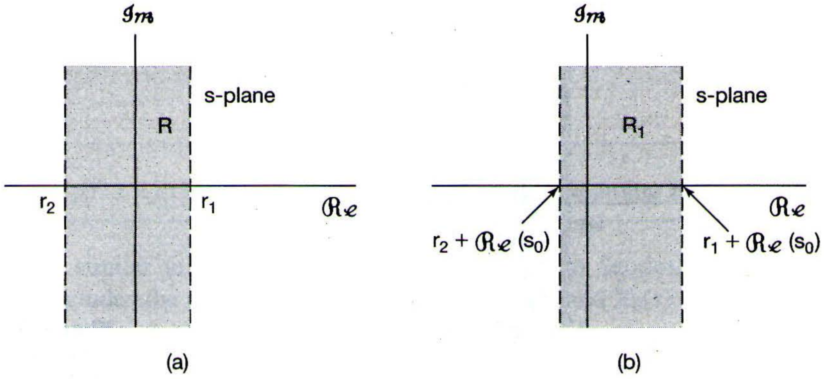
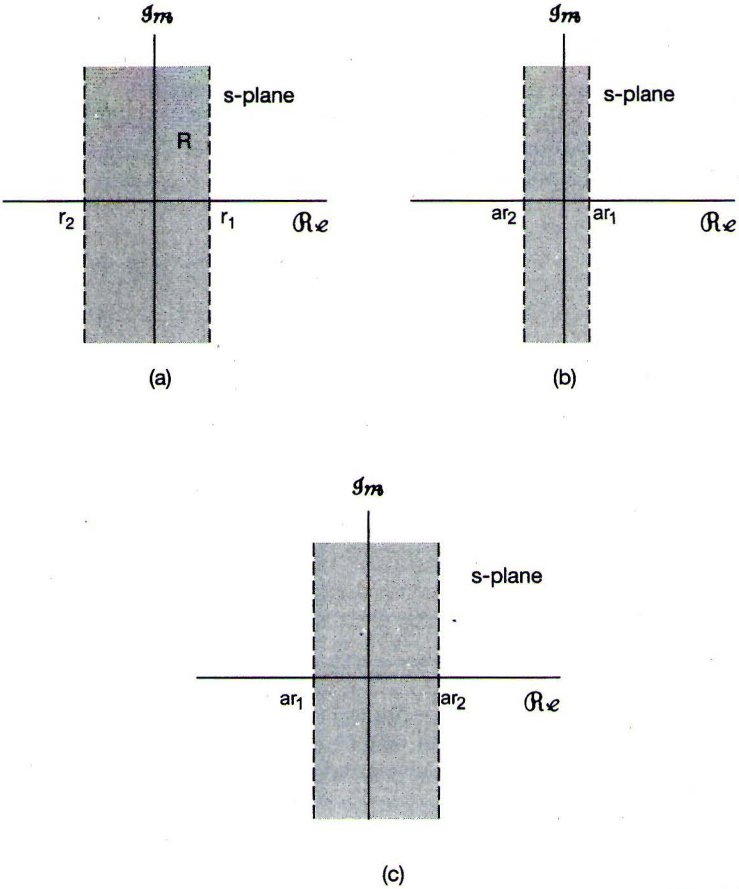


Figure 9.23 Effect on the ROC of shifting in the s-domain: (a) the ROC of $X(s)$; (b) the ROC of $X(s - s_0)$.

then

$$\boxed{x(at) \overset{\mathcal{L}}{\longleftrightarrow} \frac{1}{|a|}X\left(\frac{s}{a}\right), \quad \text{with ROC } R_1 = aR.} \tag{9.90}$$

That is, for any value s in R [which is illustrated in Figure 9.24(a)], the value as will be in R_1, as illustrated in Figure 9.24(b) for a positive value of $a < 1$. Note that, for $0 < a < 1$, there is a compression in the size of the ROC of $X(s)$ by a factor of a, as depicted in Figure 9.24(b), while for $a > 1$, the ROC is expanded by a factor of a. Also, eq. (9.90) implies that if a is negative, the ROC undergoes a reversal plus a scaling. In particular, as depicted in Figure 9.24(c), the ROC of $1/|a|X(s/a)$ for $0 > a > -1$ involves



$\mathcal{I}m$
s-plane
R
r_2
r_1
$\mathcal{R}e$
(a)
$\mathcal{I}m$
s-plane
ar_2
ar_1
$\mathcal{R}e$
(b)
$\mathcal{I}m$
s-plane
ar_1
ar_2
$\mathcal{R}e$
(c)

Figure 9.24 Effect on the ROC of time scaling: (a) ROC of $X(s)$; (b) ROC of $(1/|a|)X(s/a)$ for $0 < a < 1$; (c) ROC of $(1/|a|)X(s/a)$ for $0 > a > -1$.

a reversal about the $j\omega$-axis, together with a change in the size of the ROC by a factor of $|a|$. Thus, time reversal of $x(t)$ results in a reversal of the ROC. That is,

$$x(-t) \overset{\mathcal{L}}{\longleftrightarrow} X(-s), \quad \text{with ROC} = -R. \tag{9.91}$$

9.5.5 Conjugation

If

$$x(t) \overset{\mathcal{L}}{\longleftrightarrow} X(s), \quad \text{with ROC} = R, \tag{9.92}$$

then

$$\boxed{x^*(t) \overset{\mathcal{L}}{\longleftrightarrow} X^*(s^*), \quad \text{with ROC} = R.} \tag{9.93}$$

Therefore,

$$X(s) = X^*(s^*) \quad \text{when } x(t) \text{ is real.} \tag{9.94}$$

Consequently, if $x(t)$ is real and if $X(s)$ has a pole or zero at $s = s_0$ (i.e., if $X(s)$ is unbounded or zero at $s = s_0$), then $X(s)$ also has a pole or zero at the complex conjugate point $s = s_0^*$. For example, the transform $X(s)$ for the real signal $x(t)$ in Example 9.4 has poles at $s = 1 \pm 3j$ and zeros at $s = (-5 \pm j\sqrt{71})/2$.

9.5.6 Convolution Property

If

$$x_1(t) \overset{\mathcal{L}}{\longleftrightarrow} X_1(s), \qquad \text{with } ROC = R_1,$$

and

$$x_2(t) \overset{\mathcal{L}}{\longleftrightarrow} X_2(s), \qquad \text{with } ROC = R_2,$$

then

$$\boxed{x_1(t) * x_2(t) \overset{\mathcal{L}}{\longleftrightarrow} X_1(s)X_2(s), \quad \text{with ROC containing } R_1 \cap R_2.} \tag{9.95}$$

In a manner similar to the linearity property set forth in Section 9.5.1, the ROC of $X_1(s)X_2(s)$ includes the intersection of the ROCs of $X_1(s)$ and $X_2(s)$ and may be larger if pole-zero cancellation occurs in the product. For example, if

$$X_1(s) = \frac{s+1}{s+2}, \qquad \mathcal{Re}\{s\} > -2, \tag{9.96}$$

and

$$X_2(s) = \frac{s+2}{s+1}, \qquad \mathcal{R}e\{s\} > -1, \tag{9.97}$$

then $X_1(s)X_2(s) = 1$, and its ROC is the entire s-plane.

As we saw in Chapter 4, the convolution property in the context of the Fourier transform plays an important role in the analysis of linear time-invariant systems. In Sections 9.7 and 9.8 we will exploit in some detail the convolution property for Laplace transforms for the analysis of LTI systems in general and, more specifically, for the class of systems represented by linear constant-coefficient differential equations.

9.5.7 Differentiation in the Time Domain

If

$$x(t) \overset{\mathcal{L}}{\longleftrightarrow} X(s), \qquad \text{with ROC} = R,$$

then

$$\boxed{\frac{dx(t)}{dt} \overset{\mathcal{L}}{\longleftrightarrow} sX(s), \qquad \text{with ROC containing } R.} \tag{9.98}$$

This property follows by differentiating both sides of the inverse Laplace transform as expressed in equation (9.56). Specifically, let

$$x(t) = \frac{1}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty} X(s)e^{st}ds.$$

Then

$$\frac{dx(t)}{dt} = \frac{1}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty} sX(s)e^{st}ds. \tag{9.99}$$

Consequently, $dx(t)/dt$ is the inverse Laplace transform of $sX(s)$. The ROC of $sX(s)$ includes the ROC of $X(s)$ and may be larger if $X(s)$ has a first-order pole at $s = 0$ that is canceled by the multiplication by s. For example, if $x(t) = u(t)$, then $X(s) = 1/s$, with an ROC that is $\mathcal{R}e\{s\} > 0$. The derivative of $x(t)$ is an impulse with an associated Laplace transform that is unity and an ROC that is the entire s-plane.

9.5.8 Differentiation in the s-Domain

Differentiating both sides of the Laplace transform equation (9.3), i.e.,

$$X(s) = \int_{-\infty}^{+\infty} x(t)e^{-st}dt,$$

we obtain

$$\frac{dX(s)}{ds} = \int_{-\infty}^{+\infty} (-t)x(t)e^{-st}dt.$$

Consequently, if

$$x(t) \overset{\mathcal{L}}{\longleftrightarrow} X(s), \qquad \text{with ROC} = R,$$

then

$$\boxed{-tx(t) \overset{\mathcal{L}}{\longleftrightarrow} \frac{dX(s)}{ds}, \qquad \text{with ROC} = R.} \tag{9.100}$$

The next two examples illustrate the use of this property.

Example 9.14

Let us find the Laplace transform of

$$x(t) = te^{-at}u(t). \tag{9.101}$$

Since

$$e^{-at}u(t) \overset{\mathcal{L}}{\longleftrightarrow} \frac{1}{s+a}, \qquad \mathcal{R}e\{s\} > -a,$$

it follows from eq. (9.100) that

$$te^{-at}u(t) \overset{\mathcal{L}}{\longleftrightarrow} -\frac{d}{ds}\left[\frac{1}{s+a}\right] = \frac{1}{(s+a)^2}, \qquad \mathcal{R}e\{s\} > -a. \tag{9.102}$$

In fact, by repeated application of eq. (9.100), we obtain

$$\frac{t^2}{2}e^{-at}u(t) \overset{\mathcal{L}}{\longleftrightarrow} \frac{1}{(s+a)^3}, \qquad \mathcal{R}e\{s\} > -a, \tag{9.103}$$

and, more generally,

$$\frac{t^{n-1}}{(n-1)!}e^{-at}u(t) \overset{\mathcal{L}}{\longleftrightarrow} \frac{1}{(s+a)^n}, \qquad \mathcal{R}e\{s\} > -a. \tag{9.104}$$

As the next example illustrates, this specific Laplace transform pair is particularly useful when applying partial-fraction expansion to the determination of the inverse Laplace transform of a rational function with multiple-order poles.

Example 9.15

Consider the Laplace transform

$$X(s) = \frac{2s^2+5s+5}{(s+1)^2(s+2)}, \qquad \mathcal{R}e\{s\} > -1.$$

Applying the partial-fraction expansion method described in the appendix, we can write

$$X(s) = \frac{2}{(s+1)^2} - \frac{1}{(s+1)} + \frac{3}{s+2}, \qquad \mathcal{R}e\{s\} > -1. \tag{9.105}$$

Since the ROC is to the right of the poles at $s = -1$ and -2, the inverse transform of each of the terms is a right-sided signal, and, applying eqs. (9.14) and (9.104), we obtain the inverse transform

$$x(t) = [2te^{-t} - e^{-t} + 3e^{-2t}]u(t).$$

9.5.9 Integration in the Time Domain

If

$$x(t) \overset{\mathfrak{L}}{\longleftrightarrow} X(s), \qquad \text{with ROC} = R,$$

then

$$\boxed{\int_{-\infty}^{t} x(\tau)d\tau \overset{\mathfrak{L}}{\longleftrightarrow} \frac{1}{s}X(s), \quad \begin{array}{r}\text{with ROC containing}\\ R \cap \{\Re e\{s\} > 0\}.\end{array}} \tag{9.106}$$

This property is the inverse of the differentiation property set forth in Section 9.5.7. It can be derived using the convolution property presented in Section 9.5.6. Specifically,

$$\int_{-\infty}^{t} x(\tau)d\tau = u(t) * x(t). \tag{9.107}$$

From Example 9.1, with $a = 0$,

$$u(t) \overset{\mathfrak{L}}{\longleftrightarrow} \frac{1}{s}, \qquad \Re e\{s\} > 0, \tag{9.108}$$

and thus, from the convolution property,

$$u(t) * x(t) \overset{\mathfrak{L}}{\longleftrightarrow} \frac{1}{s}X(s), \tag{9.109}$$

with an ROC that contains the intersection of the ROC of $X(s)$ and the ROC of the Laplace transform of $u(t)$ in eq. (9.108), which results in the ROC given in eq. (9.106).

9.5.10 The Initial- and Final-Value Theorems

Under the specific constraints that $x(t) = 0$ for $t < 0$ and that $x(t)$ contains no impulses or higher order singularities at the origin, one can directly calculate, from the Laplace transform, the initial value $x(0^+)$—i.e., $x(t)$ as t approaches zero from positive values of t. Specifically the *initial-value theorem* states that

$$x(0^+) = \lim_{s\to\infty} sX(s), \tag{9.110}$$

Also, if $x(t) = 0$ for $t < 0$ and, in addition, $x(t)$ has a finite limit as $t \to \infty$, then the *final-value theorem* says that

$$\lim_{t\to\infty} x(t) = \lim_{s\to 0} sX(s). \tag{9.111}$$

The derivation of these results is considered in Problem 9.53.

Example 9.16

The initial- and final-value theorems can be useful in checking the correctness of the Laplace transform calculations for a signal. For example, consider the signal $x(t)$ in Example 9.4. From eq. (9.24), we see that $x(0^+) = 2$. Also, using eq. (9.29), we find that

$$\lim_{s \to \infty} sX(s) = \lim_{s \to \infty} \frac{2s^3 + 5s^2 + 12s}{s^3 + 4s^2 + 14s + 20} = 2,$$

which is consistent with the initial-value theorem in eq. (9.110).

9.5.11 Table of Properties

In Table 9.1, we summarize the properties developed in this section. In Section 9.7, many of these properties are used in applying the Laplace transform to the analysis and characterization of linear time-invariant systems. As we have illustrated in several examples, the various properties of Laplace transforms and their ROCs can provide us with

TABLE 9.1 PROPERTIES OF THE LAPLACE TRANSFORM

Section	Property	Signal	Laplace Transform	ROC
		$x(t)$ $x_1(t)$ $x_2(t)$	$X(s)$ $X_1(s)$ $X_2(s)$	R R_1 R_2
9.5.1	Linearity	$ax_1(t) + bx_2(t)$	$aX_1(s) + bX_2(s)$	At least $R_1 \cap R_2$
9.5.2	Time shifting	$x(t - t_0)$	$e^{-st_0}X(s)$	R
9.5.3	Shifting in the s-Domain	$e^{s_0 t}x(t)$	$X(s - s_0)$	Shifted version of R (i.e., s is in the ROC if $s - s_0$ is in R)
9.5.4	Time scaling	$x(at)$	$\frac{1}{\lvert a \rvert} X\left(\frac{s}{a}\right)$	Scaled ROC (i.e., s is in the ROC if s/a is in R)
9.5.5	Conjugation	$x^*(t)$	$X^*(s^*)$	R
9.5.6	Convolution	$x_1(t) * x_2(t)$	$X_1(s)X_2(s)$	At least $R_1 \cap R_2$
9.5.7	Differentiation in the Time Domain	$\frac{d}{dt}x(t)$	$sX(s)$	At least R
9.5.8	Differentiation in the s-Domain	$-tx(t)$	$\frac{d}{ds}X(s)$	R
9.5.9	Integration in the Time Domain	$\int_{-\infty}^{t} x(\tau)d(\tau)$	$\frac{1}{s}X(s)$	At least $R \cap \{\mathcal{R}e\{s\} > 0\}$
	Initial- and Final-Value Theorems			

9.5.10	If $x(t) = 0$ for $t < 0$ and $x(t)$ contains no impulses or higher-order singularities at $t = 0$, then
	$x(0^+) = \lim_{s \to \infty} sX(s)$
	If $x(t) = 0$ for $t < 0$ and $x(t)$ has a finite limit as $t \longrightarrow \infty$, then
	$\lim_{t \to \infty} x(t) = \lim_{s \to 0} sX(s)$			

considerable information about a signal and its transform that can be useful either in characterizing the signal or in checking a calculation. In Sections 9.7 and 9.8 and in some of the problems at the end of this chapter, we give several other examples of the uses of these properties.

9.6 SOME LAPLACE TRANSFORM PAIRS

As we indicated in Section 9.3, the inverse Laplace transform can often be easily evaluated by decomposing $X(s)$ into a linear combination of simpler terms, the inverse transform of each of which can be recognized. Listed in Table 9.2 are a number of useful Laplace

TABLE 9.2 LAPLACE TRANSFORMS OF ELEMENTARY FUNCTIONS

Transform pair	Signal	Transform	ROC
1	$\delta(t)$	1	All s
2	$u(t)$	$\frac{1}{s}$	$\Re e\{s\} > 0$
3	$-u(-t)$	$\frac{1}{s}$	$\Re e\{s\} < 0$
4	$\frac{t^{n-1}}{(n-1)!}u(t)$	$\frac{1}{s^n}$	$\Re e\{s\} > 0$
5	$-\frac{t^{n-1}}{(n-1)!}u(-t)$	$\frac{1}{s^n}$	$\Re e\{s\} < 0$
6	$e^{-\alpha t}u(t)$	$\frac{1}{s+\alpha}$	$\Re e\{s\} > -\alpha$
7	$-e^{-\alpha t}u(-t)$	$\frac{1}{s+\alpha}$	$\Re e\{s\} < -\alpha$
8	$\frac{t^{n-1}}{(n-1)!}e^{-\alpha t}u(t)$	$\frac{1}{(s+\alpha)^n}$	$\Re e\{s\} > -\alpha$
9	$-\frac{t^{n-1}}{(n-1)!}e^{-\alpha t}u(-t)$	$\frac{1}{(s+\alpha)^n}$	$\Re e\{s\} < -\alpha$
10	$\delta(t-T)$	e^{-sT}	All s
11	$[\cos \omega_0 t]u(t)$	$\frac{s}{s^2+\omega_0^2}$	$\Re e\{s\} > 0$
12	$[\sin \omega_0 t]u(t)$	$\frac{\omega_0}{s^2+\omega_0^2}$	$\Re e\{s\} > 0$
13	$[e^{-\alpha t}\cos \omega_0 t]u(t)$	$\frac{s+\alpha}{(s+\alpha)^2+\omega_0^2}$	$\Re e\{s\} > -\alpha$
14	$[e^{-\alpha t}\sin \omega_0 t]u(t)$	$\frac{\omega_0}{(s+\alpha)^2+\omega_0^2}$	$\Re e\{s\} > -\alpha$
15	$u_n(t) = \frac{d^n\delta(t)}{dt^n}$	s^n	All s
16	$u_{-n}(t) = \underbrace{u(t) * \cdots * u(t)}_{n \text{ times}}$	$\frac{1}{s^n}$	$\Re e\{s\} > 0$

transform pairs. Transform pair 1 follows directly from eq. (9.3). Transform pairs 2 and 6 follow directly from Example 9.1 with $a = 0$ and $a = \alpha$, respectively. Transform pair 4 was developed in Example 9.14 using the differentiation property. Transform pair 8 follows from transform pair 4 using the property set forth in Section 9.5.3. Transform pairs 3, 5, 7, and 9 are based on transform pairs 2, 4, 6 and 8, respectively, together with the time-scaling property of section 9.5.4 with $a = -1$. Similarly, transform pairs 10 through 16 can all be obtained from earlier ones in the table using appropriate properties in Table 9.1 (see Problem 9.55).

9.7 ANALYSIS AND CHARACTERIZATION OF LTI SYSTEMS USING THE LAPLACE TRANSFORM

One of the important applications of the Laplace transform is in the analysis and characterization of LTI systems. Its role for this class of systems stems directly from the convolution property (Section 9.5.6). Specifically, the Laplace transforms of the input and output of an LTI system are related through multiplication by the Laplace transform of the impulse response of the system. Thus,

$$Y(s) = H(s)X(s). \tag{9.112}$$

where $X(s)$, $Y(s)$, and $H(s)$ are the Laplace transforms of the input, output, and impulse response of the system, respectively. Equation (9.112) is the counterpart, in the context of Laplace transforms, of eq. (4.56) for Fourier transform. Also, from our discussion in Section 3.2 on the response of LTI systems to complex exponentials, if the input to an LTI system is $x(t) = e^{st}$, with s in the ROC of $H(s)$, then the output will be $H(s)e^{st}$; i.e., e^{st} is an eigenfunction of the system with eigenvalue equal to the Laplace transform of the impulse response.

If the ROC of $H(s)$ includes the imaginary axis, then for $s = j\omega$, $H(s)$ is the frequency response of the LTI system. In the broader context of the Laplace transform, $H(s)$ is commonly referred to as the system function or, alternatively, the *transfer function*. Many properties of LTI systems can be closely associated with the characteristics of the *system function* in the s-plane. We illustrate this next by examining several important properties and classes of systems.

9.7.1 Causality

For a causal LTI system, the impulse response is zero for $t < 0$ and thus is right sided. Consequently, from the discussion in Section 9.2, we see that

> The ROC associated with the system function for a causal system is a right-half plane.

It should be stressed, however, that the converse of this statement is not necessarily true. That is, as illustrated in Example 9.19 to follow, an ROC to the right of the rightmost pole does not guarantee that a system is causal; rather, it guarantees only that the impulse response is right sided. However, if $H(s)$ is *rational*, then, as illustrated in Examples 9.17 and 9.18 to follow, we can determine whether the system is causal simply by checking to see if its ROC is a right-half plane. Specifically,

> For a system with a rational system function, causality of the system is equivalent to the ROC being the right-half plane to the right of the rightmost pole.

Example 9.17

Consider a system with impulse response

$$h(t) = e^{-t}u(t). \tag{9.113}$$

Since $h(t) = 0$ for $t < 0$, this system is causal. Also, the system function can be obtained from Example 9.1:

$$H(s) = \frac{1}{s+1}, \qquad \mathcal{R}e\{s\} > -1. \tag{9.114}$$

In this case, the system function is rational and the ROC in eq. (9.114) is to the right of the rightmost pole, consistent with our statement that causality for systems with rational system functions is equivalent to the ROC being to the right of the rightmost pole.

Example 9.18

Consider a system with impulse response

$$h(t) = e^{-|t|}.$$

Since $h(t) \neq 0$ for $t < 0$, this system is not causal. Also, from Example 9.7, the system function is

$$H(s) = \frac{-2}{s^2-1}, \qquad -1 < \mathcal{R}e\{s\} < +1.$$

Thus, $H(s)$ is rational and has an ROC that is *not* to the right of the the rightmost pole, consistent with the fact that the system is not causal.

Example 9.19

Consider the system function

$$H(s) = \frac{e^s}{s+1}, \qquad \mathcal{R}e\{s\} > -1. \tag{9.115}$$

For this system, the ROC is to the right of the rightmost pole. Therefore, the impulse response must be right sided. To determine the impulse response, we first use the result of Example 9.1:

$$e^{-t}u(t) \overset{\mathcal{L}}{\longleftrightarrow} \frac{1}{s+1}, \qquad \mathcal{R}e\{s\} > -1. \tag{9.116}$$

Next, from the time-shifting property of Section 9.5.2 [eq. (9.87)], the factor e^s in eq. (9.115) can be accounted for by a shift in the time function in eq. (9.116). Then

$$e^{-(t+1)}u(t+1) \overset{\mathcal{L}}{\longleftrightarrow} \frac{e^s}{s+1}, \qquad \mathcal{R}e\{s\} > -1, \tag{9.117}$$

so that the impulse response associated with the system is

$$h(t) = e^{-(t+1)}u(t+1), \tag{9.118}$$

which is nonzero for $-1 < t < 0$. Hence, the system is not causal. This example serves as a reminder that causality implies that the ROC is to the right of the rightmost pole, but the converse is not in general true, unless the system function is rational.

In an exactly analogous manner, we can deal with the concept of anticausality. A system is *anticausal* if its impulse response $h(t) = 0$ for $t > 0$. Since in that case $h(t)$ would be left sided, we know from Section 9.2 that the ROC of the system function $H(s)$ would have to be a left-half plane. Again, in general, the converse is not true. That is, if the ROC of $H(s)$ is a left-half plane, all we know is that $h(t)$ is left sided. However, if $H(s)$ is rational, then having an ROC to the left of the leftmost pole is equivalent to the system being anticausal.

9.7.2 Stability

The ROC of $H(s)$ can also be related to the stability of a system. As mentioned in Section 2.3.7, the stability of an LTI system is equivalent to its impulse response being absolutely integrable, in which case (Section 4.4) the Fourier transform of the impulse response converges. Since the Fourier transform of a signal equals the Laplace transform evaluated along the $j\omega$-axis, we have the following:

> An LTI system is stable if and only if the ROC of its system function H(s) includes the entire $j\omega$-axis [i.e., $\mathcal{R}e\{s\} = 0$].

Example 9.20

Let us consider an LTI system with system function

$$H(s) = \frac{s-1}{(s+1)(s-2)}. \tag{9.119}$$

Since the ROC has not been specified, we know from our discussion in Section 9.2 that there are several different ROCs and, consequently, several different system impulse responses that can be associated with the algebraic expression for $H(s)$ given in eq. (9.119).

If, however, we have information about the causality or stability of the system, the appropriate ROC can be identified. For example, if the system is known to be *causal*, the ROC will be that indicated in Figure 9.25(a), with impulse response

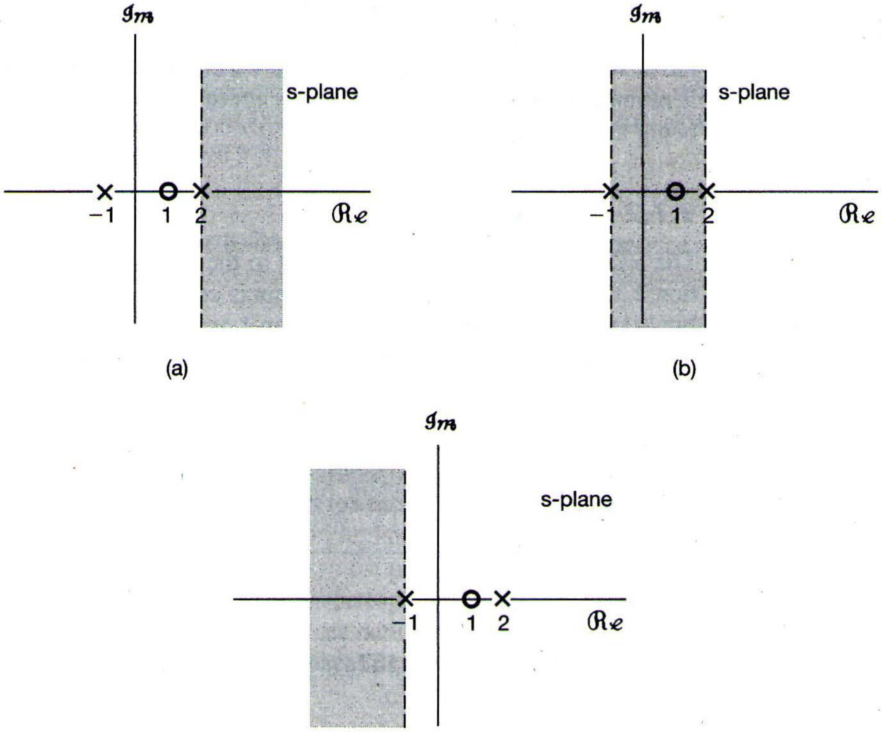
$$h(t) = \left(\frac{2}{3}e^{-t} + \frac{1}{3}e^{2t}\right)u(t). \tag{9.120}$$

Note that this particular choice of ROC does not include the $j\omega$-axis, and consequently, the corresponding system is unstable (as can be checked by observing that $h(t)$ is not absolutely integrable). On the other hand, if the system is known to be *stable*, the ROC is that given in Figure 9.25(b), and the corresponding impulse response is

$$h(t) = \frac{2}{3}e^{-t}u(t) - \frac{1}{3}e^{2t}u(-t),$$

which is absolutely integrable. Finally, for the ROC in Figure 9.25(c), the system is anticausal and unstable, with

$$h(t) = -\left(\frac{2}{3}e^{-t} + \frac{1}{3}e^{2t}\right)u(-t).$$



(a)

(b)

(c)

Figure 9.25 Possible ROCs for the system function of Example 9.20 with poles at $s = -1$ and $s = 2$ and a zero at $s = 1$: (a) causal, unstable system; (b) noncausal, stable system; (c) anticausal, unstable system.

It is perfectly possible, of course, for a system to be stable (or unstable) and have a system function that is not rational. For example, the system function in eq. (9.115) is not rational, and its impulse response in eq. (9.118) is absolutely integrable, indicating that the system is stable. However, for systems with rational system functions, stability is easily interpreted in terms of the poles of the system. For example, for the pole-zero plot in Figure 9.25, stability corresponds to the choice of an ROC that is between the two poles, so that the $j\omega$-axis is contained in the ROC.

For one particular and very important class of systems, stability can be characterized very simply in terms of the locations of the poles. Specifically, consider a causal LTI system with a rational system function $H(s)$. Since the system is causal, the ROC is to the right of the rightmost pole. Consequently, for this system to be stable (i.e., for the ROC to include the $j\omega$-axis), the rightmost pole of $H(s)$ must be to the *left* of the $j\omega$-axis. That is,

> A causal system with rational system function $H(s)$ is stable if and only if all of the poles of $H(s)$ lie in the left-half of the s-plane—i.e., all of the poles have negative real parts.

Example 9.21

Consider again the causal system in Example 9.17. The impulse response in eq. (9.113) is absolutely integrable, and thus the system is stable. Consistent with this, we see that the pole of $H(s)$ in eq. (9.114) is at $s = -1$, which is in the left-half of the s-plane. In contrast, the causal system with impulse response

$$h(t) = e^{2t}u(t)$$

is unstable, since $h(t)$ is not absolutely integrable. Also, in this case

$$H(s) = \frac{1}{s-2}, \qquad \Re e\{s\} > 2,$$

so the system has a pole at $s = 2$ in the right half of the s-plane.

Example 9.22

Let us consider the class of causal second-order systems previously discussed in Sections 9.4.2 and 6.5.2. The impulse response and system function are, respectively,

$$h(t) = M[e^{c_1 t} - e^{c_2 t}]u(t) \tag{9.121}$$

and

$$H(s) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} = \frac{\omega_n^2}{(s-c_1)(s-c_2)}, \tag{9.122}$$

where

$$c_1 = -\zeta\omega_n + \omega_n\sqrt{\zeta^2 - 1}, \tag{9.123}$$

$$c_2 = -\zeta\omega_n - \omega_n\sqrt{\zeta^2 - 1}, \tag{9.124}$$

$$M = \frac{\omega_n}{2\sqrt{\zeta^2 - 1}}. \tag{9.125}$$

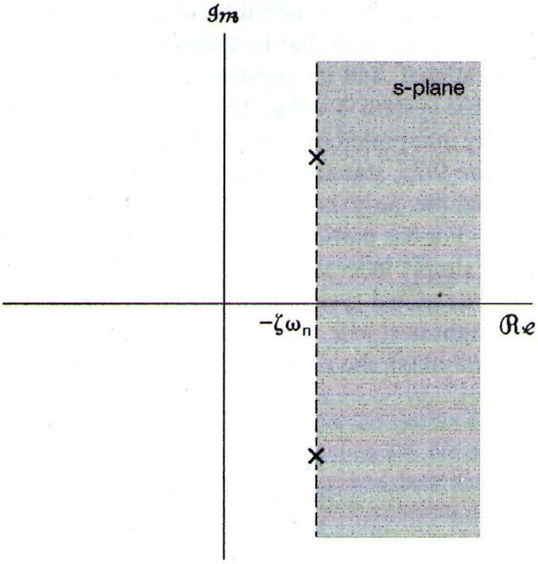


Figure 9.26 Pole locations and ROC for a causal second-order system with $\zeta < 0$.

In Figure 9.19, we illustrated the pole locations for $\zeta > 0$. In Figure 9.26, we illustrate the pole locations for $\zeta < 0$. As is evident from the latter figure and from eqs. (9.124) and (9.125), for $\zeta < 0$ both poles have positive real parts. Consequently, for $\zeta < 0$, the causal second-order system cannot be stable. This is also evident in eq. (9.121), since, with $\mathcal{R}e\{c_1\} > 0$ and $\mathcal{R}e\{c_2\} > 0$, each term grows exponentially as t increases, and thus $h(t)$ cannot be absolutely integrable.

9.7.3 LTI Systems Characterized by Linear Constant-Coefficient Differential Equations

In Section 4.7, we discussed the use of the Fourier transform to obtain the frequency response of an LTI system characterized by a linear constant-coefficient differential equation without first obtaining the impulse response or time-domain solution. In an exactly analogous manner, the properties of the Laplace transform can be exploited to directly obtain the system function for an LTI system characterized by a linear constant-coefficient differential equation. We illustrate this procedure in the next example.

Example 9.23

Consider an LTI system for which the input $x(t)$ and output $y(t)$ satisfy the linear constant-coefficient differential equation

$$\frac{dy(t)}{dt} + 3y(t) = x(t). \tag{9.126}$$

Applying the Laplace transform to both sides of eq. (9.126), and using the linearity and differentiation properties set forth in Sections 9.5.1 and 9.5.7, respectively [(eqs. (9.82) and (9.98)], we obtain the algebraic equation

$$sY(s) + 3Y(s) = X(s). \tag{9.127}$$

Since, from eq. (9.112), the system function is

$$H(s) = \frac{Y(s)}{X(s)},$$

we obtain, for this system,

$$H(s) = \frac{1}{s+3}. \tag{9.128}$$

This, then, provides the algebraic expression for the system function, but not the region of convergence. In fact, as we discussed in Section 2.4, the differential equation itself is not a complete specification of the LTI system, and there are, in general, different impulse responses, all consistent with the differential equation. If, in addition to the differential equation, we know that the system is causal, then the ROC can be inferred to be to the right of the rightmost pole, which in this case corresponds to $\mathcal{Re}\{s\} > -3$. If the system were known to be anticausal, then the ROC associated with $H(s)$ would be $\mathcal{Re}\{s\} < -3$. The corresponding impulse response in the causal case is

$$h(t) = e^{-3t}u(t), \tag{9.129}$$

whereas in the anticausal case it is

$$h(t) = -e^{-3t}u(-t). \tag{9.130}$$

The same procedure used to obtain $H(s)$ from the differential equation in Example 9.23 can be applied more generally. Consider a general linear constant-coefficient differential equation of the form

$$\sum_{k=0}^{N} a_k \frac{d^k y(t)}{dt^k} = \sum_{k=0}^{M} b_k \frac{d^k x(t)}{dt^k}. \tag{9.131}$$

Applying the Laplace transform to both sides and using the linearity and differentiation properties repeatedly, we obtain

$$\left(\sum_{k=0}^{N} a_k s^k\right) Y(s) = \left(\sum_{k=0}^{M} b_k s^k\right) X(s), \tag{9.132}$$

or

$$H(s) = \frac{\left\{\sum_{k=0}^{M} b_k s^k\right\}}{\left\{\sum_{k=0}^{N} a_k s^k\right\}}. \tag{9.133}$$

Thus, the system function for a system specified by a differential equation is always rational, with zeros at the solutions of

$$\sum_{k=0}^{M} b_k s^k = 0 \tag{9.134}$$

and poles at the solutions of

$$\sum_{k=0}^{N} a_k s^k = 0. \tag{9.135}$$

Consistently with our previous discussion, eq. (9.133) does not include a specification of the region of convergence of $H(s)$, since the linear constant-coefficient differential equation by itself does not constrain the region of convergence. However, with additional information, such as knowledge about the stability or causality of the system, the region of convergence can be inferred. For example, if we impose the condition of initial rest on the system, so that it is causal, the ROC will be to the right of the rightmost pole.

Example 9.24

An *RLC* circuit whose capacitor voltage and inductor current are initially zero constitutes an LTI system describable by a linear constant-coefficient differential equation. Consider the series *RLC* circuit in Figure 9.27. Let the voltage across the voltage source be the input signal $x(t)$, and let the voltage measured across the capacitor be the output signal $y(t)$. Equating the sum of the voltages across the resistor, inductor, and capacitor with the source voltage, we obtain

$$RC\frac{dy(t)}{dt} + LC\frac{d^2y(t)}{dt^2} + y(t) = x(t). \tag{9.136}$$

Applying eq. (9.133), we obtain

$$H(s) = \frac{1/LC}{s^2 + (R/L)s + (1/LC)}. \tag{9.137}$$

As shown in Problem 9.64, if the values of R, L, and C are all positive, the poles of this system function will have negative real parts, and consequently, the system will be stable.

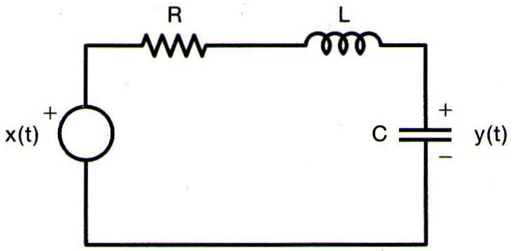


Figure 9.27 A series *RLC* circuit.

9.7.4 Examples Relating System Behavior to the System Function

As we have seen, system properties such as causality and stability can be directly related to the system function and its characteristics. In fact, each of the properties of Laplace transforms that we have described can be used in this way to relate the behavior of the system to the system function. In this section, we give several examples illustrating this.

Example 9.25

Suppose we know that if the input to an LTI system is

$$x(t) = e^{-3t}u(t),$$

then the output is

$$y(t) = [e^{-t} - e^{-2t}]u(t).$$

As we now show, from this knowledge we can determine the system function for this system and from this can immediately deduce a number of other properties of the system.

Taking Laplace transforms of $x(t)$ and $y(t)$, we get

$$X(s) = \frac{1}{s+3}, \qquad \mathcal{R}e\{s\} > -3,$$

and

$$Y(s) = \frac{1}{(s+1)(s+2)}, \qquad \mathcal{R}e\{s\} > -1.$$

From eq. (9.112), we can then conclude that

$$H(s) = \frac{Y(s)}{X(s)} = \frac{s+3}{(s+1)(s+2)} = \frac{s+3}{s^2+3s+2}.$$

Furthermore, we can also determine the ROC for this system. In particular, we know from the convolution property set forth in Section 9.5.6 that the ROC of $Y(s)$ must include at least the intersections of the ROCs of $X(s)$ and $H(s)$. Examining the three possible choices for the ROC of $H(s)$ (i.e., to the left of the pole at $s = -2$, between the poles at -2 and -1, and to the right of the pole at $s = -1$), we see that the only choice that is consistent with the ROCs of $X(s)$ and $Y(s)$ is $\mathcal{R}e\{s\} > -1$. Since this is to the right of the rightmost pole of $H(s)$, we conclude that $H(s)$ is causal, and since both poles of $H(s)$ have negative real parts, it follows that the system is stable. Moreover, from the relationship between eqs. (9.131) and (9.133), we can specify the differential equation that, together with the condition of initial rest, characterizes the system:

$$\frac{d^2y(t)}{dt^2} + 3\frac{dy(t)}{dt} + 2y(t) = \frac{dx(t)}{dt} + 3x(t).$$

Example 9.26

Suppose that we are given the following information about an LTI system:

1. The system is causal.
2. The system function is rational and has only two poles, at $s = -2$ and $s = 4$.

3. If $x(t) = 1$, then $y(t) = 0$.
4. The value of the impulse response at $t = 0^+$ is 4.

From this information we would like to determine the system function of the system.

From the first two facts, we know that the system is unstable (since it is causal and has a pole at $s = 4$ with positive real part) and that the system function is of the form

$$H(s) = \frac{p(s)}{(s+2)(s-4)} = \frac{p(s)}{s^2 - 2s - 8},$$

where $p(s)$ is a polynomial in s. Because the response $y(t)$ to the input $x(t) = 1 = e^{0 \cdot t}$ must equal $H(0) \cdot e^{0 \cdot t} = H(0)$, we conclude, from fact 3, that $p(0) = 0$—i.e., that $p(s)$ must have a root at $s = 0$ and thus is of the form

$$p(s) = sq(s),$$

where $q(s)$ is another polynomial in s.

Finally, from fact 4 and the initial-value theorem in Section 9.5.10, we see that

$$\lim_{s\to\infty} sH(s) = \lim_{s\to\infty} \frac{s^2 q(s)}{s^2 - 2s - 8} = 4. \tag{9.138}$$

As $s \to \infty$, the terms of highest power in s in both the numerator and the denominator of $sH(s)$ dominate and thus are the only ones of importance in evaluating eq. (9.138). Furthermore, if the numerator has higher degree than the denominator, the limit will diverge. Consequently, we can obtain a finite nonzero value for the limit only if the degree of the numerator of $sH(s)$ is the same as the degree of the denominator. Since the degree of the denominator is 2, we conclude that, for eq. (9.138) to hold, $q(s)$ must be a constant—i.e., $q(s) = K$. We can evaluate this constant by evaluating

$$\lim_{s\to\infty} \frac{Ks^2}{s^2 - 2s - 8} = \lim_{s\to\infty} \frac{Ks^2}{s^2} = K. \tag{9.139}$$

Equating eqs. (9.138) and (9.139), we see that $K = 4$, and thus,

$$H(s) = \frac{4s}{(s+2)(s-4)}.$$

Example 9.27

Consider a stable and causal system with impulse response $h(t)$ and system function $H(s)$. Suppose $H(s)$ is rational, contains a pole at $s = -2$, and does not have a zero at the origin. The location of all other poles and zeros is unknown. For each of the following statements let us determine whether we can definitely say that it is true, whether we can definitely say that it is false, or whether there is insufficient information to ascertain the statement's truth:

(a) $\mathfrak{F}\{h(t)e^{3t}\}$ converges.

(b) $\int_{-\infty}^{+\infty} h(t)\,dt = 0$.

(c) $th(t)$ is the impulse response of a causal and stable system.

(d) $dh(t)/dt$ contains at least one pole in its Laplace transform.

(e) $h(t)$ has finite duration.

(f) $H(s) = H(-s)$.

(g) $\lim_{s\to\infty} H(s) = 2$.

Statement (a) is false, since $\mathfrak{F}\{h(t)e^{3t}\}$ corresponds to the value of the Laplace transform of $h(t)$ at $s = -3$. If this converges, it implies that $s = -3$ is in the ROC. A causal and stable system must always have its ROC to the *right* of all of its poles. However, $s = -3$ is not to the right of the pole at $s = -2$.

Statement (b) is false, because it is equivalent to stating that $H(0) = 0$. This contradicts the fact that $H(s)$ does not have a zero at the origin.

Statement (c) is true. According to Table 9.1, the property set forth in Section 9.5.8, the Laplace transform of $th(t)$ has the same ROC as that of $H(s)$. This ROC includes the $j\omega$-axis, and therefore, the corresponding system is stable. Also, $h(t) = 0$ for $t < 0$ implies that $th(t) = 0$ for $t < 0$. Thus, $th(t)$ represents the impulse response of a causal system.

Statement (d) is true. According to Table 9.1, $dh(t)/dt$ has the Laplace transform $sH(s)$. The multiplication by s does not eliminate the pole at $s = -2$.

Statement (e) is false. If $h(t)$ is of finite duration, then if its Laplace transform has any points in its ROC, the ROC must be the entire s-plane. However, this is not consistent with $H(s)$ having a pole at $s = -2$.

Statement (f) is false. If it were true, then, since $H(s)$ has a pole at $s = -2$, it must also have a pole at $s = 2$. This is inconsistent with the fact that all the poles of a causal and stable system must be in the left half of the s-plane.

The truth of statement (g) cannot be ascertained with the information given. The statement requires that the degree of the numerator and denominator of $H(s)$ be equal, and we have insufficient information about $H(s)$ to determine whether this is the case.

9.7.5 Butterworth Filters

In Example 6.3 we briefly introduced the widely-used class of LTI systems known as Butterworth filters. The filters in this class have a number of properties, including the characteristics of the magnitude of the frequency response of each of these filters in the passband, that make them attractive for practical implementation. As a further illustration of the usefulness of Laplace transforms, in this section we use Laplace transform techniques to determine the system function of a Butterworth filter from the specification of its frequency response magnitude.

An *N*th-order lowpass Butterworth filter has a frequency response the square of whose magnitude is given by

$$|B(j\omega)|^2 = \frac{1}{1 + (j\omega/j\omega_c)^{2N}}, \tag{9.140}$$

where N is the order of the filter. From eq. (9.140), we would like to determine the system function $B(s)$ that gives rise to $|B(j\omega)|^2$. We first note that, by definition,

$$|B(j\omega)|^2 = B(j\omega)B^*(j\omega). \tag{9.141}$$

If we restrict the impulse response of the Butterworth filter to be real, then from the property of conjugate symmetry for Fourier transforms,

$$B^*(j\omega) = B(-j\omega), \tag{9.142}$$

so that

$$B(j\omega)B(-j\omega) = \frac{1}{1 + (j\omega/j\omega_c)^{2N}}. \tag{9.143}$$

Next, we note that $B(s)|_{s=j\omega} = B(j\omega)$, and consequently, from eq. (9.143),

$$B(s)B(-s) = \frac{1}{1 + (s/j\omega_c)^{2N}}. \tag{9.144}$$

The roots of the denominator polynomial corresponding to the combined poles of $B(s)B(-s)$ are at

$$s = (-1)^{1/2N}(j\omega_c). \tag{9.145}$$

Equation (9.145) is satisfied for any value $s = s_p$ for which

$$|s_p| = \omega_c \tag{9.146}$$

and

$$\sphericalangle s_p = \frac{\pi(2k+1)}{2N} + \frac{\pi}{2}, \qquad k \text{ an integer}; \tag{9.147}$$

that is,

$$s_p = \omega_c \exp\left(j\left[\frac{\pi(2k+1)}{2N} + \pi/2\right]\right). \tag{9.148}$$

In Figure 9.28 we illustrate the positions of the poles of $B(s)B(-s)$ for $N = 1, 2, 3$, and 6. In general, the following observations can be made about these poles:

1. There are $2N$ poles equally spaced in angle on a circle of radius ω_c in the s-plane.
2. A pole never lies on the $j\omega$-axis and occurs on the σ-axis for N odd, but not for N even.
3. The angular spacing between the poles of $B(s)B(-s)$ is π/N radians.

To determine the poles of $B(s)$ given the poles of $B(s)B(-s)$, we observe that the poles of $B(s)B(-s)$ occur in pairs, so that if there is a pole at $s = s_p$, then there is also a pole at $s = -s_p$. Consequently, to construct $B(s)$, we choose one pole from each pair. If we restrict the system to be stable and causal, then the poles that we associate with $B(s)$ are the poles along the semicircle in the left-half plane. The pole locations specify $B(s)$ only to within a scale factor. However, from eq. (9.144), we see that $B^2(s)|_{s=0} = 1$, or equivalently, from eq. (9.140), the scale factor is chosen so that the square of the magnitude of the frequency response has unity gain at $\omega = 0$.

To illustrate the determination of $B(s)$, let us consider the cases $N = 1$, $N = 2$, and $N = 3$. In Figure 9.28 we showed the poles of $B(s)B(-s)$, as obtained from eq. (9.148).

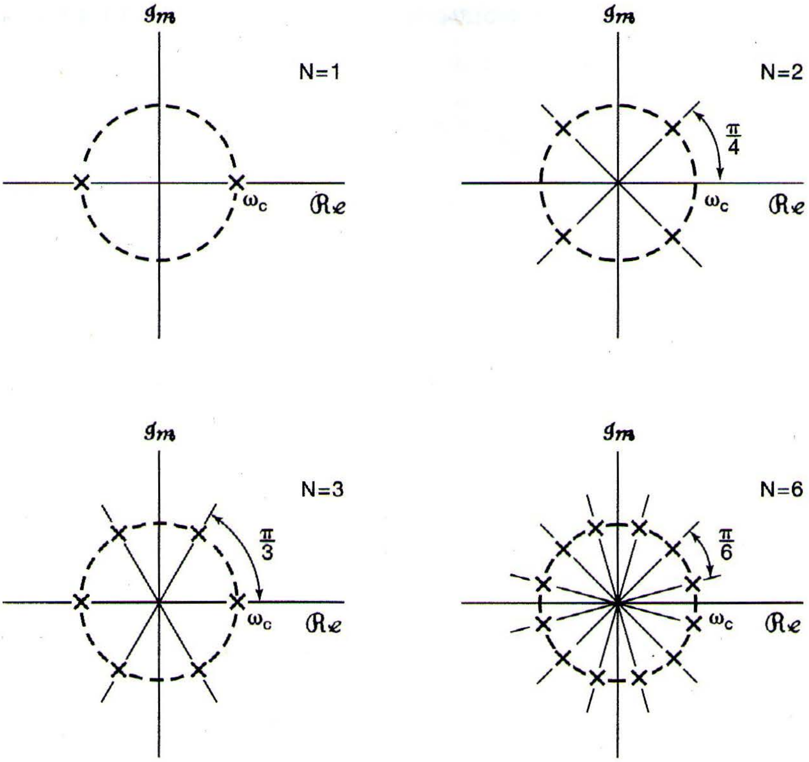

$\mathcal{Im}$
N=1
ω_c
$\mathcal{Re}$
$\mathcal{Im}$
N=2
$\frac{\pi}{4}$
ω_c
$\mathcal{Re}$
$\mathcal{Im}$
N=3
$\frac{\pi}{3}$
ω_c
$\mathcal{Re}$
$\mathcal{Im}$
N=6
$\frac{\pi}{6}$
ω_c
$\mathcal{Re}$

Figure 9.28 Position of the poles of $B(s)B(-s)$ for $N = 1, 2, 3$, and 6.

In Figure 9.29 we show the poles associated with $B(s)$ for each of these values of N. The corresponding transfer functions are:

$$N = 1: \qquad B(s) = \frac{\omega_c}{s + \omega_c}; \tag{9.149}$$

$$\begin{aligned} N = 2: \qquad B(s) &= \frac{\omega_c^2}{(s + \omega_c e^{j(\pi/4)})(s + \omega_c e^{-j(\pi/4)})} \\ &= \frac{\omega_c^2}{s^2 + \sqrt{2}\omega_c s + \omega_c^2}; \end{aligned} \tag{9.150}$$

$$\begin{aligned} N = 3: \qquad B(s) &= \frac{\omega_c^3}{(s + \omega_c)(s + \omega_c e^{j(\pi/3)})(s + \omega_c e^{-j(\pi/3)})} \\ &= \frac{\omega_c^3}{(s + \omega_c)(s^2 + \omega_c s + \omega_c^2)} \\ &= \frac{\omega_c^3}{s^3 + 2\omega_c s^2 + 2\omega_c^2 s + \omega_c^3}. \end{aligned} \tag{9.151}$$

Based on the discussion in Section 9.7.3, from $B(s)$ we can determine the associated linear constant-coefficient differential equation. Specifically, for the foregoing three values

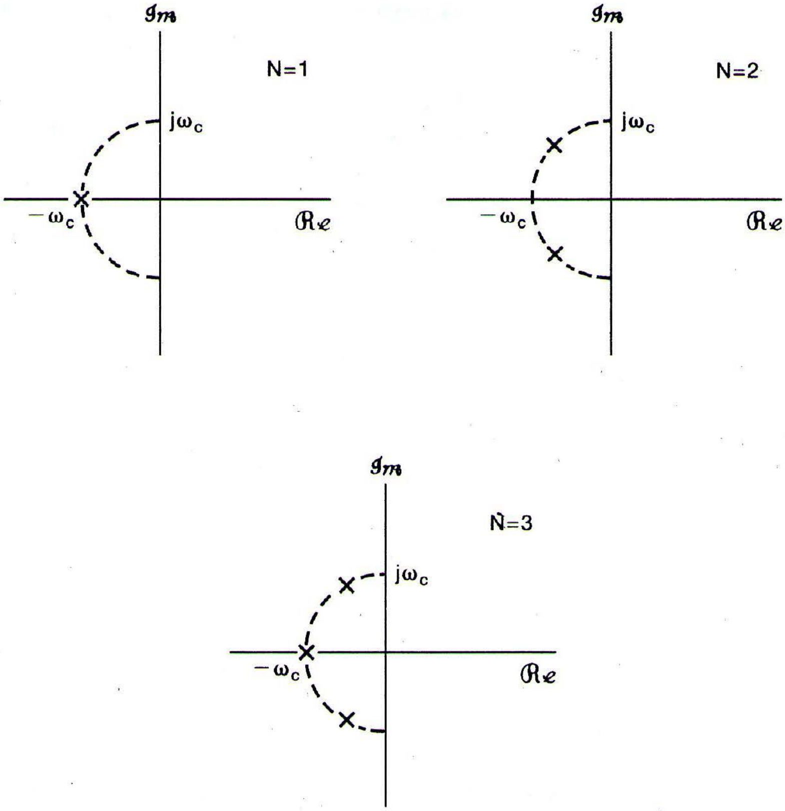


Figure 9.29 Position of the poles of $B(s)$ for $N = 1, 2,$ and 3.

of N, the corresponding differential equations are:

$$N = 1: \qquad \frac{dy(t)}{dt} + \omega_c y(t) = \omega_c x(t); \tag{9.152}$$

$$N = 2: \qquad \frac{d^2y(t)}{dt^2} + \sqrt{2}\omega_c \frac{dy(t)}{dt} + \omega_c^2 y(t) = \omega_c^2 x(t); \tag{9.153}$$

$$N = 3: \qquad \frac{d^3y(t)}{dt^3} + 2\omega_c \frac{d^2y(t)}{dt^2} + 2\omega_c^2 \frac{dy(t)}{dt} + \omega_c^3 y(t) = \omega_c^3 x(t). \tag{9.154}$$

9.8 SYSTEM FUNCTION ALGEBRA AND BLOCK DIAGRAM REPRESENTATIONS

The use of the Laplace transform allows us to replace time-domain operations such as differentiation, convolution, time shifting, and so on, with algebraic operations. We have already seen many of the benefits of this in terms of analyzing LTI systems, and in this section we take a look at another important use of system function algebra, namely, in analyzing interconnections of LTI systems and synthesizing systems as interconnections of elementary system building blocks.

9.8.1 System Functions for Interconnections of LTI Systems

Consider the parallel interconnection of two systems, as shown in Figure 9.30(a). The impulse response of the overall system is

$$h(t) = h_1(t) + h_2(t), \tag{9.155}$$

and from the linearity of the Laplace transform,

$$H(s) = H_1(s) + H_2(s). \tag{9.156}$$

Similarly, the impulse response of the series interconnection in Figure 9.30(b) is

$$h(t) = h_1(t) * h_2(t), \tag{9.157}$$

and the associated system function is

$$H(s) = H_1(s)H_2(s). \tag{9.158}$$

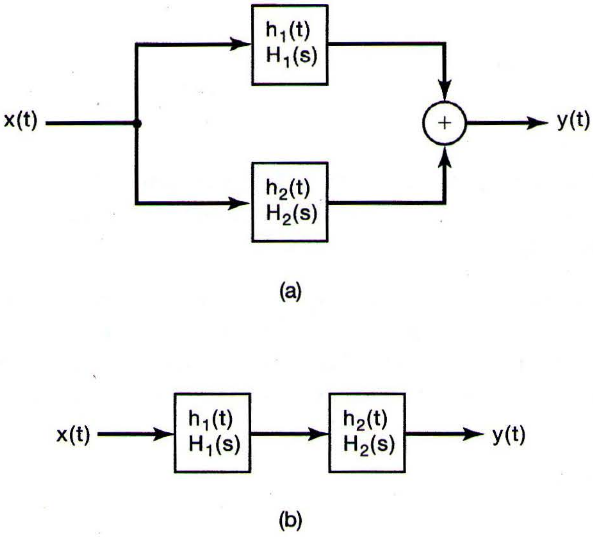


Figure 9.30 (a) Parallel interconnection of two LTI systems; (b) series combination of two LTI systems.

The utility of the Laplace transform in representing combinations of linear systems through algebraic operations extends to far more complex interconnections than the simple parallel and series combinations in Figure 9.30. To illustrate this, consider the feedback interconnection of two systems, as indicated in Figure 9.31. The design, applications, and analysis of such interconnections are treated in detail in Chapter 11. While analysis of the system in the time domain is not particularly simple, determining the overall system function from input $x(t)$ to output $y(t)$ is a straightforward algebraic manipulation. Specifically, from Figure 9.31,

$$Y(s) = H_1(s)E(s), \tag{9.159}$$

$$E(s) = X(s) - Z(s), \tag{9.160}$$

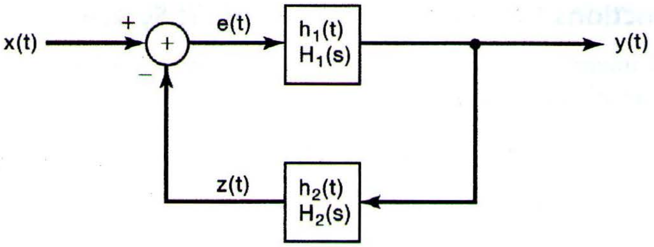


Figure 9.31 Feedback interconnection of two LTI systems.

and

$$Z(s) = H_2(s)Y(s), \tag{9.161}$$

from which we obtain the relation

$$Y(s) = H_1(s)[X(s) - H_2(s)Y(s)], \tag{9.162}$$

or

$$\frac{Y(s)}{X(s)} = H(s) = \frac{H_1(s)}{1 + H_1(s)H_2(s)}. \tag{9.163}$$

9.8.2 Block Diagram Representations for Causal LTI Systems Described by Differential Equations and Rational System Functions

In Section 2.4.3, we illustrated the block diagram representation of an LTI system described by a first-order differential equation using the basic operations of addition, multiplication by a coefficient, and integration. These same three operations can also be used to build block diagrams for higher order systems, and in this section we illustrate this in several examples.

Example 9.28

Consider the causal LTI system with system function

$$H(s) = \frac{1}{s+3}.$$

From Section 9.7.3, we know that this system can also be described by the differential equation

$$\frac{dy(t)}{dt} + 3y(t) = x(t),$$

together with the condition of initial rest. In Section 2.4.3 we constructed a block diagram representation, shown in Figure 2.32, for a first-order system such as this. An equivalent block diagram (corresponding to Figure 2.32 with $a = 3$ and $b = 1$) is shown in

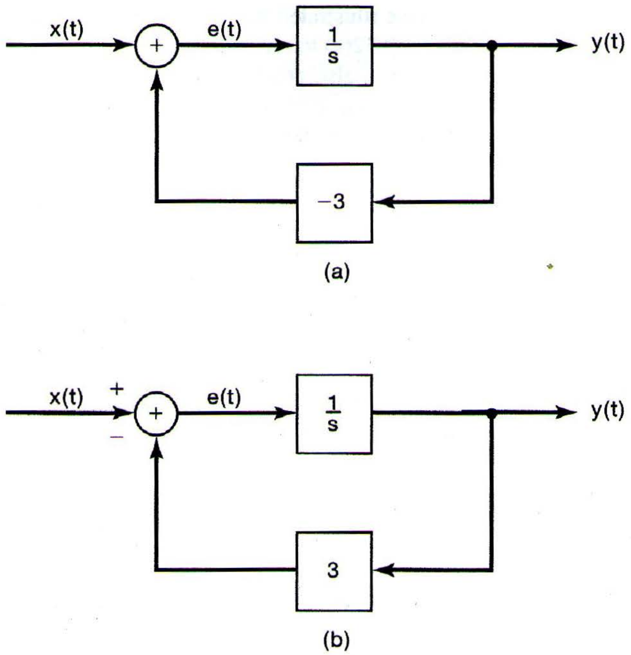


Figure 9.32 (a) Block diagram representation of the causal LTI system in Example 9.28; (b) equivalent block diagram representation.

Figure 9.32(a). Here, $1/s$ is the system function of a system with impulse response $u(t)$, i.e., it is the system function of an integrator. Also, the system function -3 in the feedback path in Figure 9.32(a) corresponds to multiplication by the coefficient -3. The block diagram in the figure involves a feedback loop much as we considered in the previous subsection and as pictured in Figure 9.31, the sole difference being that the two signals that are the inputs to the adder in Figure 9.32(a) are added, rather than subtracted as in Figure 9.31. However, as illustrated in Figure 9.32(b), by changing the sign of the coefficient in the multiplication in the feedback path, we obtain a block diagram representation of exactly the same form as Figure 9.31. Consequently, we can apply eq. (9.163) to verify that

$$H(s) = \frac{1/s}{1 + 3/s} = \frac{1}{s+3}.$$

Example 9.29

Consider now the causal LTI system with system function

$$H(s) = \frac{s+2}{s+3} = \left(\frac{1}{s+3}\right)(s+2). \tag{9.164}$$

As suggested by eq. (9.164), this system can be thought of as a cascade of a system with system function $1/(s+3)$ followed by a system with system function $s+2$, and we have illustrated this in Figure 9.33(a), in which we have used the block diagram in Figure 9.32(a) to represent $1/(s + 3)$.

It is also possible to obtain an alternative block diagram representation for the system in eq. (9.164). Using the linearity and differentiation properties of the Laplace transform, we know that $y(t)$ and $z(t)$ in Figure 9.33 (a) are related by

$$y(t) = \frac{dz(t)}{dt} + 2z(t).$$

However, the input $e(t)$ to the integrator is exactly the derivative of the output $z(t)$, so that

$$y(t) = e(t) + 2z(t),$$

which leads directly to the alternative block diagram representation shown in Figure 9.33(b). Note that the block diagram in Figure 9.33(a) requires the differentiation of $z(t)$, since

$$y(t) = \frac{dz(t)}{dt} + 2z(t)$$

In contrast, the block diagram in Figure 9.33(b) does not involve the explicit differentiation of any signal.

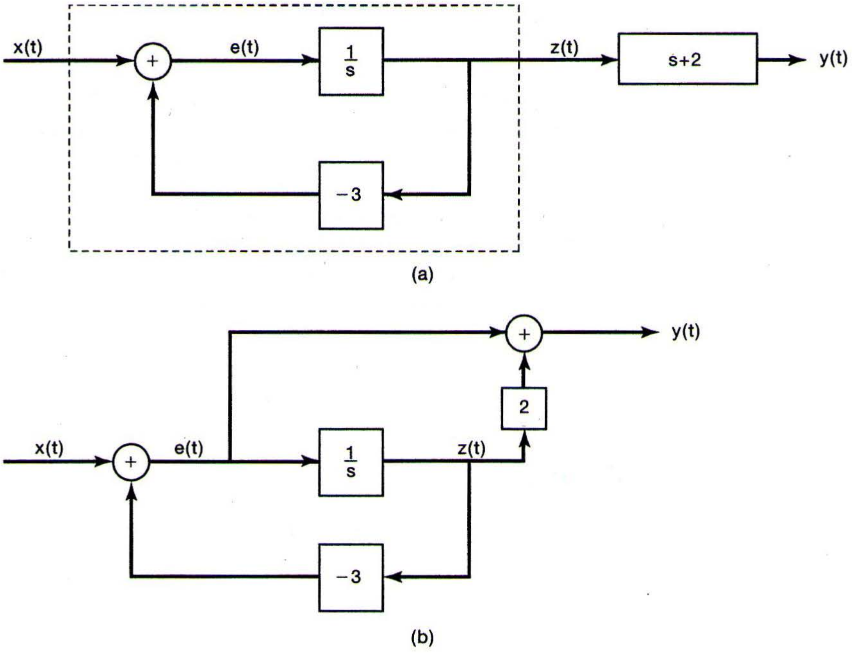


Figure 9.33 (a) Block diagram representations for the system in Example 9.29; (b) equivalent block diagram representation.

Example 9.30

Consider next a causal second-order system with system function

$$H(s) = \frac{1}{(s+1)(s+2)} = \frac{1}{s^2+3s+2}. \tag{9.165}$$

The input $x(t)$ and output $y(t)$ for this system satisfy the differential equation

$$\frac{d^2y(t)}{dt^2} + 3\frac{dy(t)}{dt} + 2y(t) = x(t). \tag{9.166}$$

By employing similar ideas to those used in the preceding examples, we obtain the block diagram representation for this system shown in Figure 9.34(a). Specifically, since the

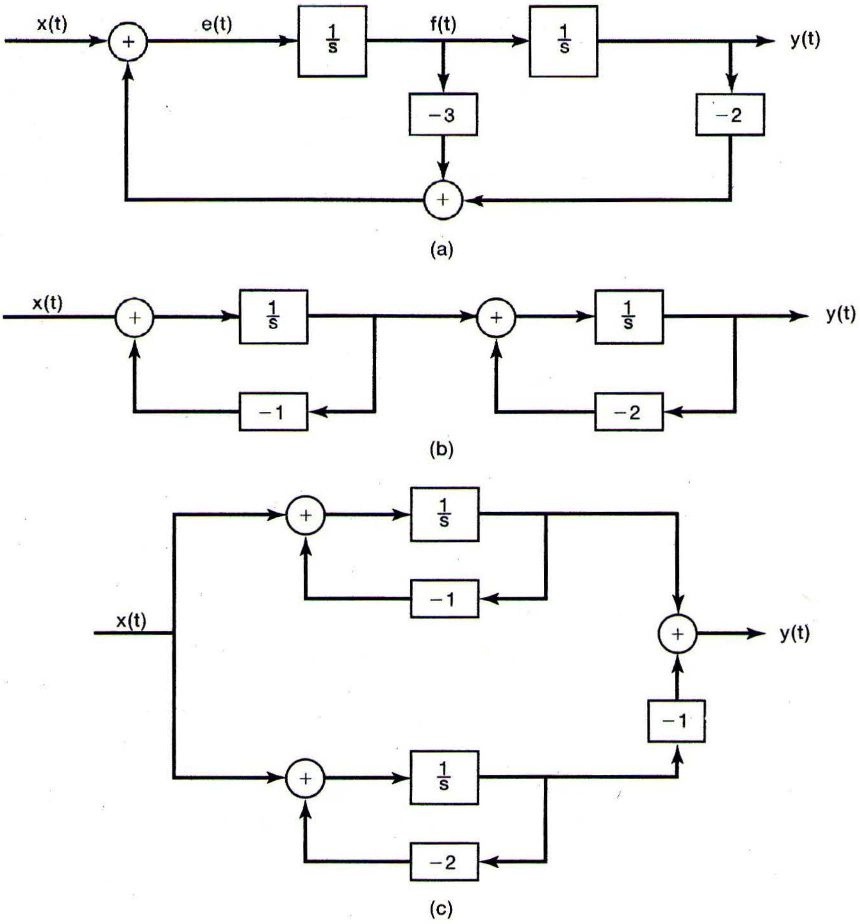


Figure 9.34 Block diagram representations for the system in Example 9.30: (a) direct form; (b) cascade form; (c) parallel form.

input to an integrator is the derivative of the output of the integrator, the signals in the block diagram are related by

$$f(t) = \frac{dy(t)}{dt},$$

$$e(t) = \frac{df(t)}{dt} = \frac{d^2y(t)}{dt}.$$

Also, eq. (9.166) can be rewritten as

$$\frac{d^2y(t)}{dt^2} = -3\frac{dy(t)}{dt} - 2y(t) + x(t),$$

or

$$e(t) = -3f(t) - 2y(t) + x(t),$$

which is exactly what is represented in Figure 9.34(a).

The block diagram in this figure is sometimes referred to as a *direct-form* representation, since the coefficients appearing in the diagram can be directly identified with the coefficients appearing in the system function or, equivalently, the differential equation. Other block diagram representations of practical importance also can be obtained after a modest amount of system function algebra. Specifically, $H(s)$ in eq. (9.165) can be rewritten as

$$H(s) = \left(\frac{1}{s+1}\right)\left(\frac{1}{s+2}\right),$$

which suggests that this system can be represented as the cascade of two first-order systems. The *cascade-form* representation corresponding to $H(s)$ is shown in Figure 9.34(b).

Alternatively, by performing a partial-fraction expansion of $H(s)$, we obtain

$$H(s) = \frac{1}{s+1} - \frac{1}{s+2},$$

which leads to the *parallel-form* representation depicted in Figure 9.34(c).

Example 9.31

As a final example, consider the system function

$$H(s) = \frac{2s^2 + 4s - 6}{s^2 + 3s + 2}. \tag{9.167}$$

Once again, using system function algebra, we can write $H(s)$ in several different forms, each of which suggests a block diagram representation. In particular, we can write

$$H(s) = \left(\frac{1}{s^2 + 3s + 2}\right)(2s^2 + 4s - 6),$$

which suggests the representation of $H(s)$ as the cascade of the system depicted in Figure 9.34(a) and the system with system function $2s^2 + 4s - 6$. However, exactly as we did in Example 9.29, we can extract the derivatives required for this second system by "tapping" the signals appearing as the inputs to the integrators in the first system. The details of this construction are examined in Problem 9.36, and the result is the direct-form block diagram shown in Figure 9.35. Once again, in the direct-form representation the coefficients appearing in the block diagram can be determined by inspection from the coefficients in the system function in eq. (9.167).

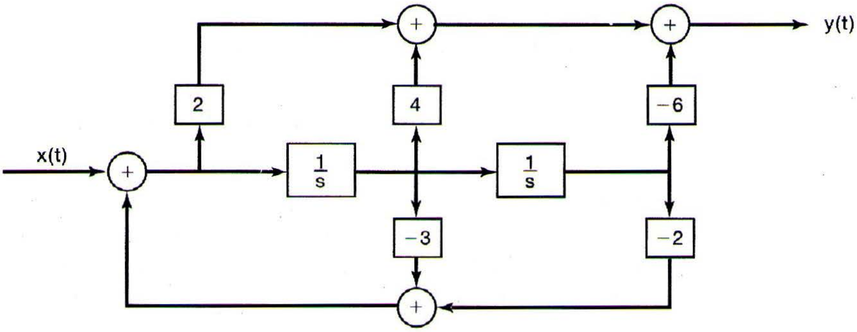


Figure 9.35 Direct-form representation for the system in Example 9.31.

Alternatively, we can write $H(s)$ in the form

$$H(s) = \left(\frac{2(s-1)}{s+2}\right)\left(\frac{s+3}{s+1}\right) \tag{9.168}$$

or

$$H(s) = 2 + \frac{6}{s+2} - \frac{8}{s+1}. \tag{9.169}$$

The first of these suggests a cascade-form representation, while the second leads to a parallel-form block diagram. These are also considered in Problem 9.36.

The methods for constructing block diagram representations for causal LTI systems described by differential equations and rational system functions can be applied equally well to higher order systems. In addition, there is often considerable flexibility in how this is done. For example, by reversing the numerators in eq. (9.168), we can write

$$H(s) = \left(\frac{s+3}{s+2}\right)\left(\frac{2(s-1)}{s+2}\right),$$

which suggests a different cascade form. Also, as illustrated in Problem 9.38, a fourth-order system function can be written as the product of two second-order system functions, each of which can be represented in a number of ways (e.g., direct form, cascade, or parallel), and it can also be written as the sum of lower order terms, each of which has several different representations. In this way, simple low-order systems can serve as building blocks for the implementation of more complex, higher order systems.

9.9 THE UNILATERAL LAPLACE TRANSFORM

In the preceding sections of this chapter, we have dealt with what is commonly called the bilateral Laplace transform. In this section, we introduce and examine a somewhat different transform, the *unilateral Laplace transform*, which is of considerable value in analyzing causal systems and, particularly, systems specified by linear constant-coefficient differential equations with nonzero initial conditions (i.e., systems that are not initially at rest).

The unilateral Laplace transform of a continuous-time signal $x(t)$ is defined as

$$\mathcal{X}(s) \triangleq \int_{0^-}^{\infty} x(t)e^{-st}\,dt, \tag{9.170}$$

where the lower limit of integration, 0^-, signifies that we include in the interval of integration any impulses or higher order singularity functions concentrated at $t = 0$. Once again we adopt a convenient shorthand notation for a signal and its unilateral Laplace transform:

$$x(t) \overset{\mathcal{UL}}{\longleftrightarrow} \mathcal{X}(s) = \mathcal{UL}\{x(t)\}. \tag{9.171}$$

Comparing eqs. (9.170) and (9.3), we see that the difference in the definitions of the unilateral and bilateral Laplace transform lies in the lower limit on the integral. The bilateral transform depends on the entire signal from $t = -\infty$ to $t = +\infty$, whereas the unilateral transform depends only on the signal from $t = 0^-$ to ∞. Consequently, two signals that differ for $t < 0$, but that are identical for $t \geq 0$, will have different bilateral Laplace transforms, but identical unilateral transforms. Similarly, any signal that is identically zero for $t < 0$ has identical bilateral and unilateral transforms.

Since the unilateral transform of $x(t)$ is identical to the bilateral transform of the signal obtained from $x(t)$ by setting its value to 0 for all $t < 0$, many of the insights, concepts, and results pertaining to bilateral transforms can be directly adapted to the unilateral case. For example, using Property 4 in Section 9.2 for right-sided signals, we see that the ROC for eq. (9.170) is always a right-half plane. The evaluation of the inverse unilateral Laplace transforms is also the same as for bilateral transforms, with the constraint that the ROC for a unilateral transform must always be a right-half plane.

9.9.1 Examples of Unilateral Laplace Transforms

To illustrate the unilateral Laplace transform, let us consider the following examples:

Example 9.32

Consider the signal

$$x(t) = \frac{t^{n-1}}{(n-1)!}e^{-at}u(t). \tag{9.172}$$

Since $x(t) = 0$ for $t < 0$, the unilateral and bilateral transforms are identical. Thus, from Table 9.2,

$$\mathcal{X}(s) = \frac{1}{(s+a)^n}, \qquad \mathcal{R}e\{s\} > -a. \tag{9.173}$$

Example 9.33

Consider next

$$x(t) = e^{-a(t+1)}u(t+1). \tag{9.174}$$

The *bilateral* transform $X(s)$ for this example can be obtained from Example 9.1 and the time-shifting property (Section 9.5.2):

$$X(s) = \frac{e^s}{s+a}, \qquad \mathcal{R}e\{s\} > -a. \tag{9.175}$$

By contrast, the unilateral transform is

$$\begin{aligned}\mathcal{X}(s) &= \int_{0^-}^{\infty} e^{-a(t+1)}u(t+1)e^{-st}\,dt \\ &= \int_{0^-}^{\infty} e^{-a}e^{-t(s+a)}\,dt \\ &= e^{-a}\frac{1}{s+a}, \qquad \mathcal{R}e\{s\} > -a.\end{aligned} \tag{9.176}$$

Thus, in this example, the unilateral and bilateral Laplace transforms are clearly different. In fact, we should recognize $\mathcal{X}(s)$ as the bilateral transform not of $x(t)$, but of $x(t)u(t)$, consistent with our earlier comment that the unilateral transform is the bilateral transform of a signal whose values for $t < 0^-$ have been set to zero.

Example 9.34

Consider the signal

$$x(t) = \delta(t) + 2u_1(t) + e^t u(t). \tag{9.177}$$

Since $x(t) = 0$ for $t < 0$, and since singularities at the origin are included in the interval of integration, the unilateral transform for $x(t)$ is the same as the bilateral transform. Specifically, using the fact (transform pair 15 in Table 9.2) that the bilateral transform of $u_n(t)$ is s^n, we have

$$\mathcal{X}(s) = X(s) = 1 + 2s + \frac{1}{s-1} = \frac{s(2s-1)}{s-1}, \qquad \mathcal{R}e\{s\} > 1. \tag{9.178}$$

Example 9.35

Consider the unilateral Laplace transform

$$\mathcal{X}(s) = \frac{1}{(s+1)(s+2)}. \tag{9.179}$$

In Example 9.9, we considered the inverse transform for a bilateral Laplace transform of the exact form as that in eq. (9.179) and for several ROCs. For the unilateral transform, the ROC must be the right-half plane to the right of the rightmost pole of $\mathcal{X}(s)$; i.e., in this case, the ROC consists of all points s with $\Re e\{s\} > -1$. We can then invert this unilateral transform exactly as in Example 9.9 to obtain

$$x(t) = [e^{-t} - e^{-2t}]u(t) \quad \text{for} \quad t > 0^{-}, \tag{9.180}$$

where we have emphasized the fact that unilateral Laplace transforms provide us with information about signals only for $t > 0^{-}$.

Example 9.36

Consider the unilateral transform

$$\mathcal{X}(s) = \frac{s^2 - 3}{s + 2}. \tag{9.181}$$

Since the degree of the numerator of $\mathcal{X}(s)$ is not strictly less than the degree of the denominator, we expand $\mathcal{X}(s)$ as

$$\mathcal{X}(s) = A + Bs + \frac{C}{s + 2}. \tag{9.182}$$

Equating eqs. (9.181) and (9.182), and clearing denominators, we obtain

$$s^2 - 3 = (A + Bs)(s + 2) + C, \tag{9.183}$$

and equating coefficients for each power of s yields

$$\mathcal{X}(s) = -2 + s + \frac{1}{s + 2}, \tag{9.184}$$

with an ROC of $\Re e\{s\} > -2$. Taking inverse transforms of each term results in

$$x(t) = -2\delta(t) + u_1(t) + e^{-2t}u(t) \quad \text{for} \quad t > 0^{-}. \tag{9.185}$$

9.9.2 Properties of the Unilateral Laplace Transform

As with the bilateral Laplace transform, the unilateral Laplace transform has a number of important properties, many of which are the same as their bilateral counterparts and several of which differ in significant ways. Table 9.3 summarizes these properties. Note that we have not included a column explicitly identifying the ROC for the unilateral Laplace transform for each signal, since the ROC of any unilateral Laplace transform is always a right-half plane. For example the ROC for a rational unilateral Laplace transform is always to the right of the rightmost pole.

Contrasting Table 9.3 with Table 9.1 for the bilateral transform, we see that, with the caveat that ROCs for unilateral Laplace transforms are always right-half planes, the linearity, s-domain shifting, time-scaling, conjugation and differentiation in the s-domain

TABLE 9.3 PROPERTIES OF THE UNILATERAL LAPLACE TRANSFORM

Property	Signal	Unilateral Laplace Transform
	$x(t)$ $x_1(t)$ $x_2(t)$	$\mathcal{X}(s)$ $\mathcal{X}_1(s)$ $\mathcal{X}_2(s)$
Linearity	$ax_1(t) + bx_2(t)$	$a\mathcal{X}_1(s) + b\mathcal{X}_2(s)$
Shifting in the s-domain	$e^{s_0 t}x(t)$	$\mathcal{X}(s - s_0)$
Time scaling	$x(at), \quad a > 0$	$\frac{1}{a}\mathcal{X}\left(\frac{s}{a}\right)$
Conjugation	$x * (t)$	$x * (s)$
Convolution (assuming that $x_1(t)$ and $x_2(t)$ are identically zero for $t < 0$)	$x_1(t) * x_2(t)$	$\mathcal{X}_1(s)\mathcal{X}_2(s)$
Differentiation in the time domain	$\frac{d}{dt}x(t)$	$s\mathcal{X}(s) - x(0^-)$
Differentiation in the s-domain	$-tx(t)$	$\frac{d}{ds}\mathcal{X}(s)$
Integration in the time domain	$\int_{0^-}^{t} x(\tau)\,d\tau$	$\frac{1}{s}\mathcal{X}(s)$

Initial- and Final-Value Theorems

If $x(t)$ contains no impulses or higher-order singularities at $t = 0$, then

$$x(0^+) = \lim_{s \to \infty} s\mathcal{X}(s)$$

$$\lim_{t \to \infty} x(t) = \lim_{s \to 0} s\mathcal{X}(s)$$

properties are identical to their bilateral counterparts. Similarly, the initial- and final-value theorems stated in Section 9.5.10 also hold for unilateral Laplace transforms.[3] The derivation of each of these properties is identical to that of its bilateral counterpart.

The convolution property for unilateral transforms also is quite similar to the corresponding property for bilateral transforms. This property states that if

$$x_1(t) = x_2(t) = 0 \quad \text{for all} \quad t < 0, \tag{9.186}$$

---

[3] In fact, the initial- and final-value theorems are basically unilateral transform properties, as they apply only to signals $x(t)$ that are identically 0 for $t < 0$.

then

$$x_1(t) * x_2(t) \overset{\mathcal{UL}}{\longleftrightarrow} \mathcal{X}_1(s)\mathcal{X}_2(s). \tag{9.187}$$

Equation (9.187) follows immediately from the bilateral convolution property, since, under the conditions of eq. (9.186), the unilateral and bilateral transforms are identical for each of the signals $x_1(t)$ and $x_2(t)$. Thus, the system analysis tools and system function algebra developed and used in this chapter apply without change to unilateral transforms, as long as we deal with causal LTI systems (for which the system function is *both* the bilateral *and* the unilateral transform of the impulse response) with inputs that are identically zero for $t < 0$. An example of this is the integration property in Table 9.3: If $x(t) = 0$ for $t < 0$, then

$$\int_{0^-}^{t} x(\tau)\,d\tau = x(t) * u(t) \overset{\mathcal{UL}}{\longleftrightarrow} \mathcal{X}(s)\mathcal{U}(s) = \frac{1}{s}\mathcal{X}(s) \tag{9.188}$$

As a second case in point, consider the following example:

Example 9.37

Suppose a causal LTI system is described by the differential equation

$$\frac{d^2y(t)}{dt^2} + 3\frac{dy(t)}{dt} + 2y(t) = x(t), \tag{9.189}$$

together with the condition of initial rest. Using eq. (9.133), we find that the system function for this system is

$$\mathcal{H}(s) = \frac{1}{s^2 + 3s + 2}. \tag{9.190}$$

Let the input to this system be $x(t) = \alpha u(t)$. In this case, the unilateral (and bilateral) Laplace transform of the output $y(t)$ is

$$\begin{aligned} \mathcal{Y}(s) &= \mathcal{H}(s)\mathcal{X}(s) = \frac{\alpha}{s(s+1)(s+2)} \\ &= \frac{\alpha/2}{s} - \frac{\alpha}{s+1} + \frac{\alpha/2}{s+2}. \end{aligned} \tag{9.191}$$

Applying Example 9.32 to each term of eq. (9.191) yields

$$y(t) = \alpha\left[\frac{1}{2} - e^{-t} + \frac{1}{2}e^{-2t}\right]u(t). \tag{9.192}$$

It is important to note that the convolution property for unilateral Laplace transforms applies only if the signals $x_1(t)$ and $x_2(t)$ in eq. (9.187) are both zero for $t < 0$. That is, while we have seen that the bilateral Laplace transform of $x_1(t) * x_2(t)$ always equals the product of the bilateral transforms of $x_1(t)$ and $x_2(t)$, the unilateral transform of $x_1(t)*x_2(t)$ in general does *not* equal the product of the unilateral transforms if either $x_1(t)$ or $x_2(t)$ is nonzero for $t < 0$. (See, for example, Problem 9.39).

A particularly important difference between the properties of the unilateral and bilateral transforms is the differentiation property. Consider a signal $x(t)$ with unilateral Laplace transform $\mathcal{X}(s)$. Then, integrating by parts, we find that the unilateral transform of $dx(t)/dt$ is given by

$$\begin{aligned}\int_{0^-}^{\infty} \frac{dx(t)}{dt} e^{-st}\,dt &= x(t)e^{-st}\Big|_{0^-}^{\infty} + s\int_{0^-}^{\infty} x(t)e^{-st}\,dt \\ &= s\mathcal{X}(s) - x(0^-).\end{aligned} \tag{9.193}$$

Similarly, a second application of this would yield the unilateral Laplace transform of $d^2x(t)/dt^2$, i.e.,

$$s^2\mathcal{X}(s) - sx(0^-) - x'(0^-), \tag{9.194}$$

where $x'(0^-)$ denotes the derivative of $x(t)$ evaluated at $t = 0^-$. Clearly, we can continue the procedure to obtain the unilateral transform of higher derivatives.

9.9.3 Solving Differential Equations Using the Unilateral Laplace Transform

A primary use of the unilateral Laplace transform is in obtaining the solution of linear constant-coefficient differential equations with nonzero initial conditions. We illustrate this in the following example:

Example 9.38

Consider the system characterized by the differential equation (9.189) with initial conditions

$$y(0^-) = \beta, \qquad y'(0^-) = \gamma. \tag{9.195}$$

Let $x(t) = \alpha u(t)$. Then, applying the unilateral transform to both sides of eq. (9.189), we obtain

$$s^2\mathcal{Y}(s) - \beta s - \gamma + 3s\mathcal{Y}(s) - 3\beta + 2\mathcal{Y}(s) = \frac{\alpha}{s}, \tag{9.196}$$

or

$$\mathcal{Y}(s) = \frac{\beta(s+3)}{(s+1)(s+2)} + \frac{\gamma}{(s+1)(s+2)} + \frac{\alpha}{s(s+1)(s+2)}, \tag{9.197}$$

where $\mathcal{Y}(s)$ is the unilateral Laplace transform of $y(t)$.

Referring to Example 9.37 and, in particular, to eq. (9.191), we see that the last term on the right-hand side of eq. (9.197) is precisely the unilateral Laplace transform of the response of the system when the initial conditions in eq. (9.195) are both zero ($\beta = \gamma = 0$). That is, the last term represents the response of the causal LTI system described by eq. (9.189) and the condition of initial rest. This response is often referred to as the *zero-state response*—i.e., the response when the initial state (the set of initial conditions in eq. (9.195)) is zero.

An analogous interpretation applies to the first two terms on the right-hand side of eq. (9.197). These terms represent the unilateral transform of the response of the system when the input is zero ($\alpha = 0$). This response is commonly referred to as the *zero-input response*. Note that the zero-input response is a linear function of the values of the initial conditions (e.g., doubling the values of both β and γ doubles the zero-input response). Furthermore, eq. (9.197) illustrates an important fact about the solution of linear constant-coefficient differential equations with nonzero initial conditions, namely, that the overall response is simply the superposition of the zero-state and the zero-input responses. The zero-state response is the response obtained by setting the initial conditions to zero—i.e., it is the response of an LTI system defined by the differential equation and the condition of initial rest. The zero-input response is the response to the initial conditions with the input set to zero. Other examples illustrating this can be found in Problems 9.20, 9.40, and 9.66.

Finally, for any values of α, β, and γ, we can, of course, expand $\mathcal{Y}(s)$ in eq. (9.197) in a partial-fraction expansion and invert to obtain $y(t)$. For example, if $\alpha = 2$, $\beta = 3$, and $\gamma = -5$, then performing a partial-fraction expansion for eq. (9.197) we find that

$$\mathcal{Y}(s) = \frac{1}{s} - \frac{1}{s+1} + \frac{3}{s+2}. \tag{9.198}$$

Applying Example 9.32 to each term then yields

$$y(t) = [1 - e^{-t} + 3e^{-2t}]u(t) \quad \text{for} \quad t > 0. \tag{9.199}$$

9.10 SUMMARY

In this chapter, we have developed and studied the Laplace transform, which can be viewed as a generalization of the Fourier transform. It is particularly useful as an analytical tool in the analysis and study of LTI systems. Because of the properties of Laplace transforms, LTI systems, including those represented by linear constant-coefficient differential equations, can be characterized and analyzed in the transform domain by algebraic manipulations. In addition, system function algebra provides a convenient tool both for analyzing interconnections of LTI systems and for constructing block diagram representations of LTI systems described by differential equations.

For signals and systems with rational Laplace transforms, the transform is often conveniently represented in the complex plane (s-plane) by marking the locations of the poles and zeros and indicating the region of convergence. From the pole-zero plot, the Fourier transform can be geometrically obtained, within a scaling factor. Causality, stability, and other characteristics are also easily identified from knowledge of the pole locations and the region of convergence.

In this chapter, we have been concerned primarily with the bilateral Laplace transform. However, we also introduced a somewhat different form of the Laplace transform known as the unilateral Laplace transform. The unilateral transform can be interpreted as the bilateral transform of a signal whose values prior to $t = 0^-$ have been set to zero. This form of the Laplace transform is especially useful for analyzing systems described by linear constant-coefficient differential equations with nonzero initial conditions.

Chapter 9 Problems

The first section of problems belongs to the basic category, and the answers are provided in the back of the book. The remaining three sections contain problems belonging to the basic, advanced, and extension categories, respectively.

BASIC PROBLEMS WITH ANSWERS

9.1. For each of the following integrals, specify the values of the real parameter σ which ensure that the integral converges:

(a) $\int_0^{\infty} e^{-5t} e^{-(\sigma+j\omega)t}\, dt$ **(b)** $\int_{-\infty}^{0} e^{-5t} e^{-(\sigma+j\omega)t}\, dt$

(c) $\int_{-5}^{5} e^{-5t} e^{-(\sigma+j\omega)t}\, dt$ **(d)** $\int_{-\infty}^{\infty} e^{-5t} e^{-(\sigma+j\omega)t}\, dt$

(e) $\int_{-\infty}^{\infty} e^{-5|t|} e^{-(\sigma+j\omega)t}\, dt$ **(f)** $\int_{-\infty}^{0} e^{-5|t|} e^{-(\sigma+j\omega)t}\, dt$

9.2. Consider the signal

$$x(t) = e^{-5t} u(t-1),$$

and denote its Laplace transform by $X(s)$.

(a) Using eq. (9.3), evaluate $X(s)$ and specify its region of convergence.

(b) Determine the values of the finite numbers A and t_0 such that the Laplace transform $G(s)$ of

$$g(t) = Ae^{-5t} u(-t - t_0)$$

has the same algebraic form as $X(s)$. What is the region of convergence corresponding to $G(s)$?

9.3. Consider the signal

$$x(t) = e^{-5t} u(t) + e^{-\beta t} u(t),$$

and denote its Laplace transform by $X(s)$. What are the constraints placed on the real and imaginary parts of β if the region of convergence of $X(s)$ is $\Re e\{s\} > -3$?

9.4. For the Laplace transform of

$$x(t) = \begin{cases} e^{t} \sin 2t, & t \leq 0 \\ 0, & t > 0 \end{cases},$$

indicate the location of its poles and its region of convergence.

9.5. For each of the following algebraic expressions for the Laplace transform of a signal, determine the number of zeros located in the finite s-plane and the number of zeros located at infinity:

(a) $\dfrac{1}{s+1} + \dfrac{1}{s+3}$

(b) $\dfrac{s+1}{s^2-1}$

(c) $\dfrac{s^3-1}{s^2+s+1}$

9.6. An absolutely integrable signal $x(t)$ is known to have a pole at $s = 2$. Answer the following questions:

(a) Could $x(t)$ be of finite duration?
(b) Could $x(t)$ be left sided?
(c) Could $x(t)$ be right sided?
(d) Could $x(t)$ be two sided?

9.7. How many signals have a Laplace transform that may be expressed as

$$\frac{(s-1)}{(s+2)(s+3)(s^2+s+1)}$$

in its region of convergence?

9.8. Let $x(t)$ be a signal that has a rational Laplace transform with exactly two poles, located at $s = -1$ and $s = -3$. If $g(t) = e^{2t}x(t)$ and $G(j\omega)$ [the Fourier transform of $g(t)$] converges, determine whether $x(t)$ is left sided, right sided, or two sided.

9.9. Given that

$$e^{-at}u(t) \overset{\mathcal{L}}{\longleftrightarrow} \frac{1}{s+a}, \qquad \mathcal{R}e\{s\} > \mathcal{R}e\{-a\},$$

determine the inverse Laplace transform of

$$X(s) = \frac{2(s+2)}{s^2+7s+12}, \qquad \mathcal{R}e\{s\} > -3.$$

9.10. Using geometric evaluation of the magnitude of the Fourier transform from the corresponding pole-zero plot, determine, for each of the following Laplace transforms, whether the magnitude of the corresponding Fourier transform is approximately lowpass, highpass, or bandpass:

(a) $H_1(s) = \dfrac{1}{(s+1)(s+3)}, \quad \mathcal{R}e\{s\} > -1$

(b) $H_2(s) = \dfrac{s}{s^2+s+1}, \quad \mathcal{R}e\{s\} > -\frac{1}{2}$

(c) $H_3(s) = \dfrac{s^2}{s^2+2s+1}, \quad \mathcal{R}e\{s\} > -1$

9.11. Use geometric evaluation from the pole-zero plot to determine the magnitude of the Fourier transform of the signal whose Laplace transform is specified as

$$X(s) = \frac{s^2-s+1}{s^2+s+1}, \qquad \mathcal{R}e\{s\} > -\frac{1}{2}.$$

9.12. Suppose we are given the following three facts about the signal $x(t)$:

1. $x(t) = 0$ for $t < 0$.
2. $x(k/80) = 0$ for $k = 1, 2, 3, \ldots$.
3. $x(1/160) = e^{-120}$.

Let $X(s)$ denote the Laplace transform of $x(t)$, and determine which of the following statements is consistent with the given information about $x(t)$:

(a) $X(s)$ has only one pole in the finite s-plane.
(b) $X(s)$ has only two poles in the finite s-plane.
(c) $X(s)$ has more than two poles in the finite s-plane.

9.13. Let

$$g(t) = x(t) + \alpha x(-t),$$

where

$$x(t) = \beta e^{-t}u(t)$$

and the Laplace transform of $g(t)$ is

$$G(s) = \frac{s}{s^2 - 1}, \qquad -1 < \mathcal{R}e\{s\} < 1.$$

Determine the values of the constants α and β.

9.14. Suppose the following facts are given about the signal $x(t)$ with Laplace transform $X(s)$:

1. $x(t)$ is real and even.
2. $X(s)$ has four poles and no zeros in the finite s-plane.
3. $X(s)$ has a pole at $s = (1/2)e^{j\pi/4}$.
4. $\int_{-\infty}^{\infty} x(t)\,dt = 4$.

Determine $X(s)$ and its ROC.

9.15. Consider two right-sided signals $x(t)$ and $y(t)$ related through the differential equations

$$\frac{dx(t)}{dt} = -2y(t) + \delta(t)$$

and

$$\frac{dy(t)}{dt} = 2x(t).$$

Determine $Y(s)$ and $X(s)$, along with their regions of convergence.

9.16. A causal LTI system S with impulse response $h(t)$ has its input $x(t)$ and output $y(t)$ related through a linear constant-coefficient differential equation of the form

$$\frac{d^3y(t)}{dt^3} + (1+\alpha)\frac{d^2y(t)}{dt^2} + \alpha(\alpha+1)\frac{dy(t)}{dt} + \alpha^2 y(t) = x(t).$$

(a) If

$$g(t) = \frac{dh(t)}{dt} + h(t),$$

how many poles does $G(s)$ have?

(b) For what real values of the parameter α is S guaranteed to be stable?

9.17. A causal LTI system S has the block diagram representation shown in Figure P9.17. Determine a differential equation relating the input $x(t)$ to the output $y(t)$ of this system.

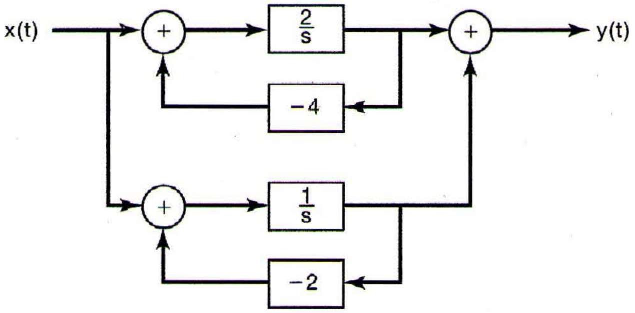


Figure P9.17

9.18. Consider the causal LTI system represented by the *RLC* circuit examined in Problem 3.20.

(a) Determine $H(s)$ and specify its region of convergence. Your answer should be consistent with the fact that the system is causal and stable.

(b) Using the pole-zero plot of $H(s)$ and geometric evaluation of the magnitude of the Fourier transform, determine whether the magnitude of the corresponding Fourier transform has an approximately lowpass, highpass, or bandpass characteristic.

(c) If the value of R is now changed to $10^{-3}\ \Omega$, determine $H(s)$ and specify its region of convergence.

(d) Using the pole-zero plot of $H(s)$ obtained in part (c) and geometric evaluation of the magnitude of the Fourier transform, determine whether the magnitude of the corresponding Fourier transform has an approximately lowpass, highpass, or bandpass characteristic.

9.19. Determine the unilateral Laplace transform of each of the following signals, and specify the corresponding regions of convergence:

(a) $x(t) = e^{-2t}u(t+1)$

(b) $x(t) = \delta(t+1) + \delta(t) + e^{-2(t+3)}u(t+1)$

(c) $x(t) = e^{-2t}u(t) + e^{-4t}u(t)$

9.20. Consider the *RL* circuit of Problem 3.19.

(a) Determine the zero-state response of this circuit when the input current is $x(t) = e^{-2t}u(t)$.

(b) Determine the zero-input response of the circuit for $t > 0^-$, given that

$$y(0^-) = 1.$$

(c) Determine the output of the circuit when the input current is $x(t) = e^{-2t}u(t)$ and the initial condition is the same as the one specified in part (b).

BASIC PROBLEMS

9.21. Determine the Laplace transform and the associated region of convergence and pole-zero plot for each of the following functions of time:

(a) $x(t) = e^{-2t}u(t) + e^{-3t}u(t)$ **(b)** $x(t) = e^{-4t}u(t) + e^{-5t}(\sin 5t)u(t)$

(c) $x(t) = e^{2t}u(-t) + e^{3t}u(-t)$ **(d)** $x(t) = te^{-2|t|}$

(e) $x(t) = |t|e^{-2|t|}$ **(f)** $x(t) = |t|e^{2t}u(-t)$

(g) $x(t) = \begin{cases} 1, & 0 \leq t \leq 1 \\ 0, & \text{elsewhere} \end{cases}$ **(h)** $x(t) = \begin{cases} t, & 0 \leq t \leq 1 \\ 2-t, & 1 \leq t \leq 2 \end{cases}$

(i) $x(t) = \delta(t) + u(t)$ **(j)** $x(t) = \delta(3t) + u(3t)$

9.22. Determine the function of time, $x(t)$, for each of the following Laplace transforms and their associated regions of convergence:

(a) $\frac{1}{s^2+9}$, $\mathcal{R}e\{s\} > 0$

(b) $\frac{s}{s^2+9}$, $\mathcal{R}e\{s\} < 0$

(c) $\frac{s+1}{(s+1)^2+9}$, $\mathcal{R}e\{s\} < -1$

(d) $\frac{s+2}{s^2+7s+12}$, $-4 < \mathcal{R}e\{s\} < -3$

(e) $\frac{s+1}{s^2+5s+6}$, $-3 < \mathcal{R}e\{s\} < -2$

(f) $\frac{(s+1)^2}{s^2-s+1}$, $\mathcal{R}e\{s\} > \frac{1}{2}$

(g) $\frac{s^2-s+1}{(s+1)^2}$, $\mathcal{R}e\{s\} > -1$

9.23. For each of the following statements about $x(t)$, and for each of the four pole-zero plots in Figure P9.23, determine the corresponding constraint on the ROC:

1. $x(t)e^{-3t}$ is absolutely integrable.
2. $x(t) * (e^{-t}u(t))$ is absolutely integrable.
3. $x(t) = 0,\ t > 1$.
4. $x(t) = 0,\ t < -1$.

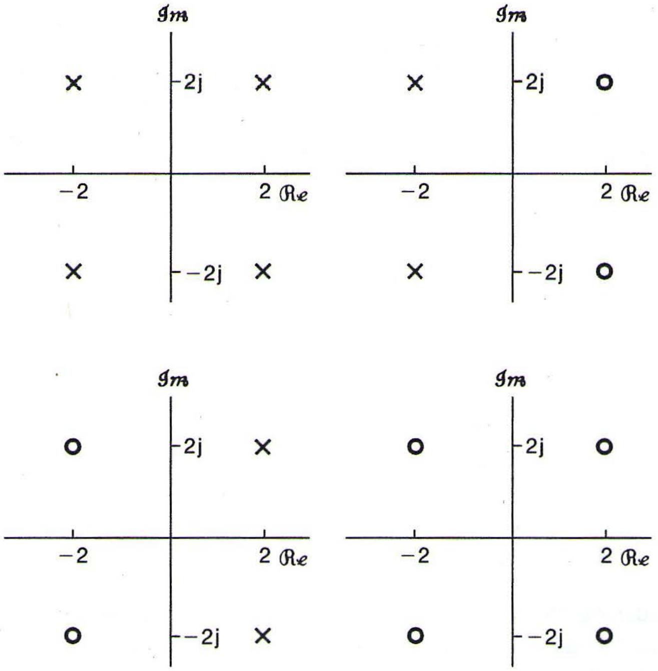


Figure P9.23

9.24. Throughout this problem, we will consider the region of convergence of the Laplace transforms always to include the $j\omega$-axis.

(a) Consider a signal $x(t)$ with Fourier transform $X(j\omega)$ and Laplace transform $X(s) = s + 1/2$. Draw the pole-zero plot for $X(s)$. Also, draw the vector whose length represents $|X(j\omega)|$ and whose angle with respect to the real axis represents $\sphericalangle X(j\omega)$ for a given ω.

(b) By examining the pole-zero plot and vector diagram in part (a), determine a different Laplace transform $X_1(s)$ corresponding to the function of time, $x_1(t)$, so that

$$|X_1(j\omega)| = |X(j\omega)|,$$

but

$$x_1(t) \neq x(t).$$

Show the pole-zero plot and associated vectors that represent $X_1(j\omega)$.

(c) For your answer in part (b), determine, again by examining the related vector diagrams, the relationship between $\sphericalangle X(j\omega)$ and $\sphericalangle X_1(j\omega)$.

(d) Determine a Laplace transform $X_2(s)$ such that

$$\sphericalangle X_2(j\omega) = \sphericalangle X(j\omega),$$

but $x_2(t)$ is not proportional to $x(t)$. Show the pole-zero plot for $X_2(s)$ and the associated vectors that represent $X_2(j\omega)$.

(e) For your answer in part (d), determine the relationship between $|X_2(j\omega)|$ and $|X(j\omega)|$.

(f) Consider a signal $x(t)$ with Laplace transform $X(s)$ for which the pole-zero plot is shown in Figure P9.24. Determine $X_1(s)$ such that $|X(j\omega)| = |X_1(j\omega)|$ and all poles and zeros of $X_1(s)$ are in the left-half of the s-plane [i.e., $\mathcal{R}e\{s\} < 0$]. Also, determine $X_2(s)$ such that $\sphericalangle X(j\omega) = \sphericalangle X_2(j\omega)$ and all poles and zeros of $X_2(s)$ are in the left-half of the s-plane.

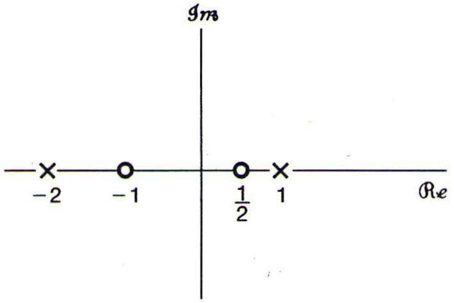
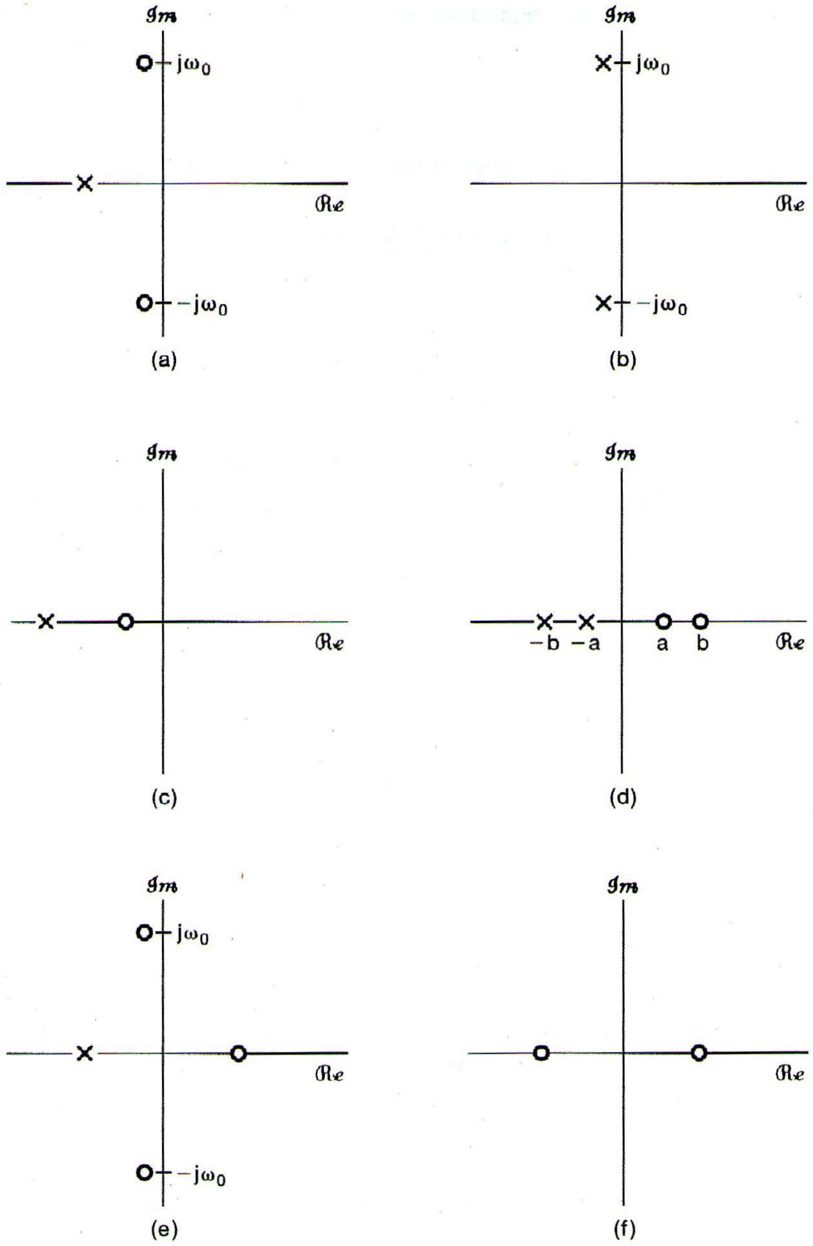


Figure P9.24

9.25. By considering the geometric determination of the Fourier transform, as developed in Section 9.4, sketch, for each of the pole-zero plots in Figure P9.25, the magnitude of the associated Fourier transform.



$\mathcal{I}m$
$j\omega_0$
$\mathcal{R}e$
$-j\omega_0$
(a)
$\mathcal{I}m$
$j\omega_0$
$\mathcal{R}e$
$-j\omega_0$
(b)
$\mathcal{I}m$
$\mathcal{R}e$
(c)
$\mathcal{I}m$
$-b$
$-a$
a
b
$\mathcal{R}e$
(d)
$\mathcal{I}m$
$j\omega_0$
$\mathcal{R}e$
$-j\omega_0$
(e)
$\mathcal{I}m$
$\mathcal{R}e$
(f)

Figure P9.25

9.26. Consider a signal $y(t)$ which is related to two signals $x_1(t)$ and $x_2(t)$ by

$$y(t) = x_1(t-2) * x_2(-t+3)$$

where

$$x_1(t) = e^{-2t}u(t) \qquad \text{and} \qquad x_2(t) = e^{-3t}u(t).$$

Given that

$$e^{-at}u(t) \overset{\mathcal{L}}{\longleftrightarrow} \frac{1}{s+a}, \quad \mathcal{R}e\{s\} > -a,$$

use properties of the Laplace transform to determine the Laplace transform $Y(s)$ of $y(t)$.

9.27. We are given the following five facts about a real signal $x(t)$ with Laplace transform $X(s)$:

1. $X(s)$ has exactly two poles.
2. $X(s)$ has no zeros in the finite s-plane.
3. $X(s)$ has a pole at $s = -1 + j$.
4. $e^{2t}x(t)$ is not absolutely integrable.
5. $X(0) = 8$.

Determine $X(s)$ and specify its region of convergence.

9.28. Consider an LTI system for which the system function $H(s)$ has the pole-zero pattern shown in Figure P9.28.

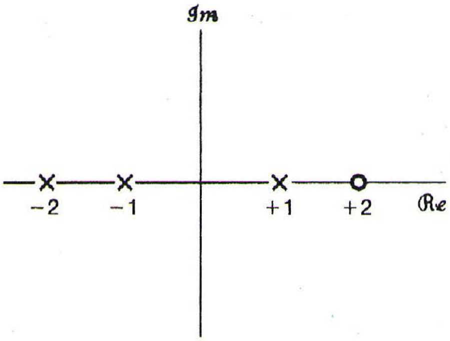


Figure P9.28

(a) Indicate all possible ROCs that can be associated with this pole-zero pattern.
(b) For each ROC identified in part (a), specify whether the associated system is stable and/or causal.

9.29. Consider an LTI system with input $x(t) = e^{-t}u(t)$ and impulse response $h(t) = e^{-2t}u(t)$.

(a) Determine the Laplace transforms of $x(t)$ and $h(t)$.
(b) Using the convolution property, determine the Laplace transform $Y(s)$ of the output $y(t)$.
(c) From the Laplace transform of $y(t)$ as obtained in part (b), determine $y(t)$.
(d) Verify your result in part (c) by explicitly convolving $x(t)$ and $h(t)$.

9.30. A pressure gauge that can be modeled as an LTI system has a time response to a unit step input given by $(1 - e^{-t} - te^{-t})u(t)$. For a certain input $x(t)$, the output is observed to be $(2 - 3e^{-t} + e^{-3t})u(t)$.

For this observed measurement, determine the true pressure input to the gauge as a function of time.

9.31. Consider a continuous-time LTI system for which the input $x(t)$ and output $y(t)$ are related by the differential equation

$$\frac{d^2y(t)}{dt^2} - \frac{dy(t)}{dt} - 2y(t) = x(t).$$

Let $X(s)$ and $Y(s)$ denote Laplace transforms of $x(t)$ and $y(t)$, respectively, and let $H(s)$ denote the Laplace transform of $h(t)$, the system impulse response.

(a) Determine $H(s)$ as a ratio of two polynomials in s. Sketch the pole-zero pattern of $H(s)$.

(b) Determine $h(t)$ for each of the following cases:
1. The system is stable.
2. The system is causal.
3. The system is *neither* stable *nor* causal.

9.32. A causal LTI system with impulse response $h(t)$ has the following properties:
1. When the input to the system is $x(t) = e^{2t}$ for all t, the output is $y(t) = (1/6)e^{2t}$ for all t.
2. The impulse response $h(t)$ satisfies the differential equation

$$\frac{dh(t)}{dt} + 2h(t) = (e^{-4t})u(t) + bu(t),$$

where b is an unknown constant.

Determine the system function $H(s)$ of the system, consistent with the information above. There should be no unknown constants in your answer; that is, the constant b should *not* appear in the answer.

9.33. The system function of a causal LTI system is

$$H(s) = \frac{s+1}{s^2+2s+2}.$$

Determine and sketch the response $y(t)$ when the input is

$$x(t) = e^{-|t|}, \qquad -\infty < t < \infty.$$

9.34. Suppose we are given the following information about a causal and stable LTI system S with impulse response $h(t)$ and a rational system function $H(s)$:
1. $H(1) = 0.2$.
2. When the input is $u(t)$, the output is absolutely integrable.
3. When the input is $tu(t)$, the output is not absolutely integrable.
4. The signal $d^2h(t)/dt^2 + 2\,dh(t)/dt + 2h(t)$ is of finite duration.
5. $H(s)$ has exactly one zero at infinity.

Determine $H(s)$ and its region of convergence.

9.35. The input $x(t)$ and output $y(t)$ of a causal LTI system are related through the block-diagram representation shown in Figure P9.35.

(a) Determine a differential equation relating $y(t)$ and $x(t)$.

(b) Is this system stable?

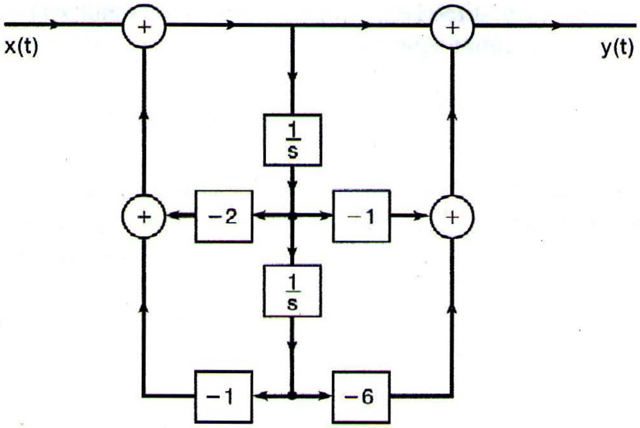


Figure P9.35

9.36. In this problem, we consider the construction of various types of block diagram representations for a causal LTI system S with input $x(t)$, output $y(t)$, and system function

$$H(s) = \frac{2s^2 + 4s - 6}{s^2 + 3s + 2}.$$

To derive the direct-form block diagram representation of S, we first consider a causal LTI system S_1 that has the same input $x(t)$ as S, but whose system function is

$$H_1(s) = \frac{1}{s^2 + 3s + 2}.$$

With the output of S_1 denoted by $y_1(t)$, the direct-form block diagram representation of S_1 is shown in Figure P9.36. The signals $e(t)$ and $f(t)$ indicated in the figure represent respective inputs into the two integrators.

(a) Express $y(t)$ (the output of S) as a linear combination of $y_1(t)$, $dy_1(t)/dt$, and $d^2y_1(t)/dt^2$.

(b) How is $dy_1(t)/dt$ related to $f(t)$?

(c) How is $d^2y_1(t)/dt^2$ related to $e(t)$?

(d) Express $y(t)$ as a linear combination of $e(t)$, $f(t)$, and $y_1(t)$.

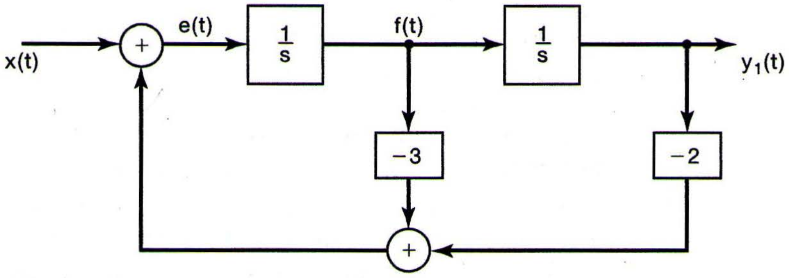


Figure P9.36

(e) Use the result from the previous part to extend the direct-form block diagram representation of S_1 and create a block diagram representation of S.

(f) Observing that

$$H(s) = \left(\frac{2(s-1)}{s+2}\right)\left(\frac{s+3}{s+1}\right),$$

draw a block diagram representation for S as a cascade combination of two subsystems.

(g) Observing that

$$H(s) = 2 + \frac{6}{s+2} - \frac{8}{s+1},$$

draw a block-diagram representation for S as a parallel combination of three subsystems.

9.37. Draw a direct-form representation for the causal LTI systems with the following system functions:

(a) $H_1(s) = \frac{s+1}{s^2+5s+6}$ **(b)** $H_2(s) = \frac{s^2-5s+6}{s^2+7s+10}$ **(c)** $H_3(s) = \frac{s}{(s+2)^2}$

9.38. Consider a fourth-order causal LTI system S whose system function is specified as

$$H(s) = \frac{1}{(s^2-s+1)(s^2+2s+1)}.$$

(a) Show that a block diagram representation for S consisting of a cascade of four first-order sections will contain multiplications by coefficients that are not purely real.

(b) Draw a block diagram representation for S as a *cascade* interconnection of two second-order systems, each of which is represented in direct form. There should be no multiplications by nonreal coefficients in the resulting block diagram.

(c) Draw a block diagram representation for S as a *parallel* interconnection of two second-order systems, each of which is represented in direct form. There should be no multiplications by nonreal coefficients in the resulting block diagram.

9.39. Let

$$x_1(t) = e^{-2t}u(t) \quad \text{and} \quad x_2(t) = e^{-3(t+1)}u(t+1).$$

(a) Determine the unilateral Laplace transform $\mathcal{X}_1(s)$ and the bilateral Laplace transform $X_1(s)$ for the signal $x_1(t)$.

(b) Determine the unilateral Laplace transform $\mathcal{X}_2(s)$ and the bilateral Laplace transform $X_2(s)$ for the signal $x_2(t)$.

(c) Take the inverse bilateral Laplace transform of the product $X_1(s)X_2(s)$ to determine the signal $g(t) = x_1(t) * x_2(t)$.

(d) Show that the inverse unilateral Laplace transform of the product $\mathcal{X}_1(s)\mathcal{X}_2(s)$ is not the same as $g(t)$ for $t > 0^-$.

9.40. Consider the system S characterized by the differential equation

$$\frac{d^3y(t)}{dt^3} + 6\frac{d^2y(t)}{dt^2} + 11\frac{dy(t)}{dt} + 6y(t) = x(t).$$

(a) Determine the zero-state response of this system for the input $x(t) = e^{-4t}u(t)$.

(b) Determine the zero-input response of the system for $t > 0^-$, given that

$$y(0^-) = 1, \qquad \left.\frac{dy(t)}{dt}\right|_{t=0^-} = -1, \qquad \left.\frac{d^2y(t)}{dt^2}\right|_{t=0^-} = 1.$$

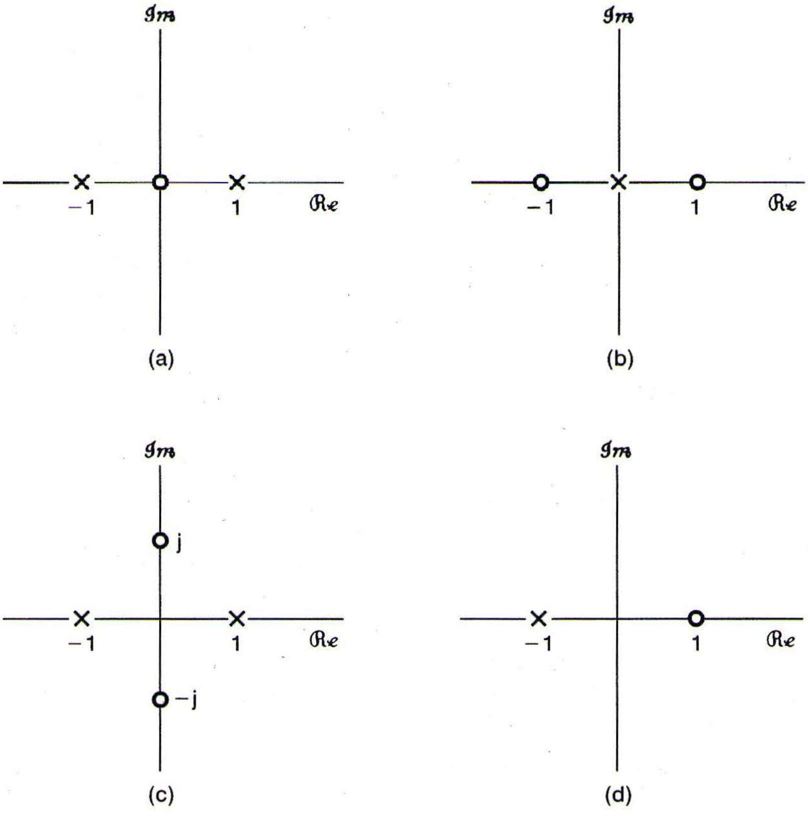
(c) Determine the output of S when the input is $x(t) = e^{-4t}u(t)$ and the initial conditions are the same as those specified in part (b).

ADVANCED PROBLEMS

9.41. **(a)** Show that, if $x(t)$ is an even function, so that $x(t) = x(-t)$, then $X(s) = X(-s)$.

(b) Show that, if $x(t)$ is an odd function, so that $x(t) = -x(-t)$, then $X(s) = -X(-s)$.

(c) Determine which, if any, of the pole-zero plots in Figure P9.41 could correspond to an even function of time. For those that could, indicate the required ROC.



(a) (b) (c) (d)

Figure P9.41

9.42. Determine whether each of the following statements is true or false. If a statement is true, construct a convincing argument for it. If it is false, give a counterexample.

(a) The Laplace transform of $t^2u(t)$ does not converge anywhere on the s-plane.

(b) The Laplace transform of $e^{t^2}u(t)$ does not converge anywhere on the s-plane.

(c) The Laplace transform of $e^{j\omega_0 t}$ does not converge anywhere on the s-plane.

(d) The Laplace transform of $e^{j\omega_0 t}u(t)$ does not converge anywhere on the s-plane.

(e) The Laplace transform of $|t|$ does not converge anywhere on the s-plane.

9.43. Let $h(t)$ be the impulse response of a causal and stable LTI system with a rational system function.

(a) Is the system with impulse response $dh(t)/dt$ guaranteed to be causal and stable?

(b) Is the system with impulse response $\int_{-\infty}^{t} h(\tau)\,d\tau$ guaranteed to be causal and unstable?

9.44. Let $x(t)$ be the sampled signal specified as

$$x(t) = \sum_{n=0}^{\infty} e^{-nT}\,\delta(t - nT),$$

where $T > 0$.

(a) Determine $X(s)$, including its region of convergence.

(b) Sketch the pole-zero plot for $X(s)$.

(c) Use geometric interpretation of the pole-zero plot to argue that $X(j\omega)$ is periodic.

9.45. Consider the LTI system shown in Figure P9.45(a) for which we are given the following information:

$$X(s) = \frac{s+2}{s-2},$$

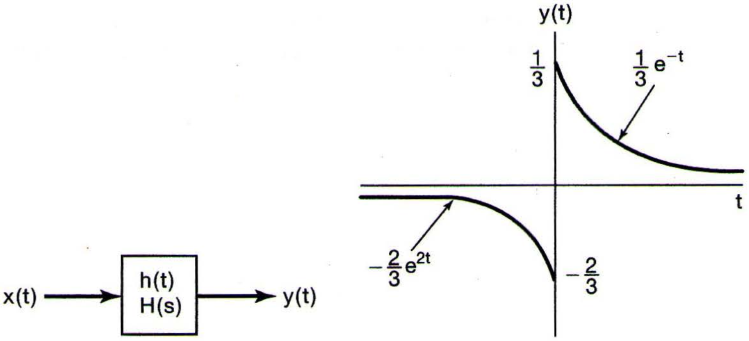
$$x(t) = 0, \qquad t > 0,$$

and

$$y(t) = -\frac{2}{3}e^{2t}u(-t) + \frac{1}{3}e^{-t}u(t). \qquad \text{[See Figure P9.45(b).]}$$

(a) Determine $H(s)$ and its region of convergence.

(b) Determine $h(t)$.



(a) (b)

Figure P9.45

(c) Using the system function $H(s)$ found in part (a), determine the output $y(t)$ if the input is

$$x(t) = e^{3t}, \qquad -\infty < t < +\infty.$$

9.46. Let $H(s)$ represent the system function for a causal, stable system. The input to the system consists of the sum of three terms, one of which is an impulse $\delta(t)$ and another a complex exponential of the form $e^{s_0 t}$, where s_0 is a complex constant. The output is

$$y(t) = -6e^{-t}u(t) + \frac{4}{34}e^{4t}\cos 3t + \frac{18}{34}e^{4t}\sin 3t + \delta(t).$$

Determine $H(s)$, consistently with this information.

9.47. The signal

$$y(t) = e^{-2t}u(t)$$

is the output of a causal all-pass system for which the system function is

$$H(s) = \frac{s-1}{s+1}.$$

(a) Find and sketch at least two possible inputs $x(t)$ that could produce $y(t)$.

(b) What is the input $x(t)$ if it is known that

$$\int_{-\infty}^{\infty} |x(t)|\, dt < \infty?$$

(c) What is the input $x(t)$ if it is known that a stable (but not necessarily causal) system exists that will have $x(t)$ as an output if $y(t)$ is the input? Find the impulse response $h(t)$ of this filter, and show by direct convolution that it has the property claimed [i.e., that $y(t) * h(t) = x(t)$].

9.48. The *inverse* of an LTI system $H(s)$ is defined as a system that, when cascaded with $H(s)$, results in an overall transfer function of unity or, equivalently, an overall impulse response that is an impulse.

(a) If $H_1(s)$ denotes the transfer function of an inverse system for $H(s)$, determine the general algebraic relationship between $H(s)$ and $H_1(s)$.

(b) Shown in Figure P9.48 is the pole-zero plot for a stable, causal system $H(s)$. Determine the pole-zero plot for the associated inverse system.

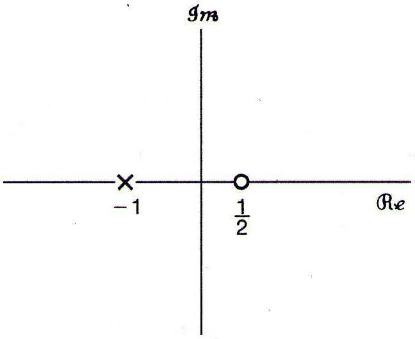


Figure P9.48

9.49. A class of systems, referred to as minimum-delay or minimum-phase systems, is sometimes defined through the statement that these systems are causal and stable and that the inverse systems are also causal and stable.

Based on the preceding definition, develop an argument to demonstrate that all poles *and* zeros of the transfer function of a minimum-delay system must be in the left half of the s-plane [i.e., $\mathcal{R}e\{s\} < 0$].

9.50. Determine whether or not each of the following statements about LTI systems is true. If a statement is true, construct a convincing argument for it. If it is false, give a counterexample.

(a) A stable continuous-time system must have all its poles in the left half of the s-plane [i.e., $\mathcal{R}e\{s\} < 0$].

(b) If a system function has more poles than zeros, and the system is causal, the step response will be continuous at $t = 0$.

(c) If a system function has more poles than zeros, and the system is not restricted to be causal, the step response can be discontinuous at $t = 0$.

(d) A stable, causal system must have all its poles and zeros in the left half of the s-plane.

9.51. Consider a stable and causal system with a real impulse response $h(t)$ and system function $H(s)$. It is known that $H(s)$ is rational, one of its poles is at $-1 + j$, one of its zeros is at $3 + j$, and it has exactly two zeros at infinity. For each of the following statements, determine whether it is true, whether it is false, or whether there is insufficient information to determine the statement's truth.

(a) $h(t)e^{-3t}$ is absolutely integrable.

(b) The ROC for $H(s)$ is $\mathcal{R}e\{s\} > -1$.

(c) The differential equation relating inputs $x(t)$ and outputs $y(t)$ for S may be written in a form having only real coefficients.

(d) $\lim_{s \to \infty} H(s) = 1$.

(e) $H(s)$ does not have fewer than four poles.

(f) $H(j\omega) = 0$ for at least one finite value of ω.

(g) If the input to S is $e^{3t} \sin t$, the output is $e^{3t} \cos t$.

9.52. As indicated in Section 9.5, many of the properties of the Laplace transform and their derivation are analogous to corresponding properties of the Fourier transform and their derivation, as developed in Chapter 4. In this problem, you are asked to outline the derivation of a number of the Laplace transform properties.

Observing the derivation for the corresponding property in Chapter 4 for the Fourier transform, derive each of the following Laplace transform properties. Your derivation must include a consideration of the region of convergence.

(a) Time shifting (Section 9.5.2)

(b) Shifting in the s-domain (Section 9.5.3)

(c) Time scaling (Section 9.5.4)

(d) Convolution property (Section 9.5.6)

9.53. As presented in Section 9.5.10, the initial-value theorem states that, for a signal $x(t)$ with Laplace transform $X(s)$ and for which $x(t) = 0$ for $t < 0$, the initial value of $x(t)$ [i.e., $x(0+)$] can be obtained from $X(s)$ through the relation

$$x(0+) = \lim_{s \to \infty} sX(s). \qquad \text{[eq. (9.110)]}$$

First, we note that, since $x(t) = 0$ for $t < 0$, $x(t) = x(t)u(t)$. Next, expanding $x(t)$ as a Taylor series at $t = 0+$, we obtain

$$x(t) = \left[x(0+) + x^{(1)}(0+)t + \cdots + x^{(n)}(0+)\frac{t^n}{n!} + \cdots\right]u(t), \qquad \text{(P9.53–1)}$$

where $x^{(n)}(0+)$ denotes the nth derivative of $x(t)$ evaluated at $t = 0+$.

(a) Determine the Laplace transform of an arbitrary term $x^{(n)}(0+)(t^n/n!)u(t)$ on the right-hand side of eq. (P9.53–1). (You may find it helpful to review Example 9.14.)

(b) From your result in part (a) and the expansion in eq. (P9.53–1), show that $X(s)$ can be expressed as

$$X(s) = \sum_{n=0}^{\infty} x^{(n)}(0+)\frac{1}{s^{n+1}}.$$

(c) Demonstrate that eq. (9.110) follows from the result of part (b).

(d) By first determining $x(t)$, verify the initial-value theorem for each of the following examples:

(1) $X(s) = \frac{1}{s+2}$

(2) $X(s) = \frac{s+1}{(s+2)(s+3)}$

(e) A more general form of the initial-value theorem states that if $x^{(n)}(0+) = 0$ for $n < N$, then $x^{(N)}(0+) = \lim_{s\to\infty} s^{N+1}X(s)$. Demonstrate that this more general statement also follows from the result in part (b).

9.54. Consider a real-valued signal $x(t)$ with Laplace transform $X(s)$.

(a) By applying complex conjugation to both sides of eq. (9.56), show that $X(s) = X^*(s^*)$.

(b) From your result in (a), show that if $X(s)$ has a pole (zero) at $s = s_0$, it must also have a pole (zero) at $s = s_0^*$; i.e., for $x(t)$ real, the poles and zeros of $X(s)$ that are not on the real axis must occur in complex conjugate pairs.

9.55. In Section 9.6, Table 9.2, we listed a number of Laplace transform pairs, and we indicated specifically how transform pairs 1 through 9 follow from Examples 9.1 and 9.14, together with various properties from Table 9.1.

By exploiting properties from Table 9.1, show how transform pairs 10 through 16 follow from transform pairs 1 through 9 in Table 9.2.

9.56. The Laplace transform is said to exist for a specific complex s if the magnitude of the transform is finite—that is, if $|X(s)| < \infty$.

Show that a *sufficient* condition for the existence of the transform $X(s)$ at $s = s_0 = \sigma_0 + j\omega_0$ is that

$$\int_{-\infty}^{+\infty} |x(t)|e^{-\sigma_0 t}\,dt < \infty.$$

In other words, show that $x(t)$ exponentially weighted by $e^{-\sigma_0 t}$ is absolutely integrable. You will need to use the result that, for a complex function $f(t)$,

$$\left| \int_a^b f(t)\,dt \right| \leq \int_a^b |f(t)|\,dt. \qquad \text{(P9.56–1)}$$

Without rigorously proving eq. (P9.56–1), argue its plausibility.

9.57. The Laplace transform $X(s)$ of a signal $x(t)$ has four poles and an unknown number of zeros. The signal $x(t)$ is known to have an impulse at $t = 0$. Determine what information, if any, this provides about the number of zeros and their locations.

9.58. Let $h(t)$ be the impulse response of a causal and stable LTI system with rational system function $H(s)$. Show that $g(t) = \mathcal{R}e\{h(t)\}$ is also the impulse response of a causal and stable system.

9.59. If $\mathcal{X}(s)$ denotes the unilateral Laplace transform of $x(t)$, determine, in terms of $\mathcal{X}(s)$, the unilateral Laplace transform of:

(a) $x(t-1)$ **(b)** $x(t+1)$

(c) $\int_{-\infty}^{\infty} x(\tau)\,d\tau$ **(d)** $\frac{d^3x(t)}{dt^3}$

EXTENSION PROBLEMS

9.60. In long-distance telephone communication, an echo is sometimes encountered due to the transmitted signal being reflected at the receiver, sent back down the line, reflected again at the transmitter, and returned to the receiver. The impulse response for a system that models this effect is shown in Figure P9.60, where we have assumed that only one echo is received. The parameter T corresponds to the one-way travel time along the communication channel, and the parameter α represents the attenuation in amplitude between transmitter and receiver.

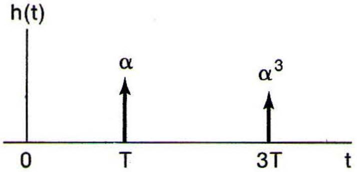


Figure P9.60

(a) Determine the system function $H(s)$ and associated region of convergence for the system.

(b) From your result in part (a), you should observe that $H(s)$ does not consist of a ratio of polynomials. Nevertheless, it is useful to represent it in terms of poles and zeros, where, as usual, the zeros are the values of s for which $H(s) = 0$

and the poles are the values of s for which $1/H(s) = 0$. For the system function found in part (a), determine the zeros and demonstrate that there are no poles.

(c) From your result in part (b), sketch the pole-zero plot for $H(s)$.

(d) By considering the appropriate vectors in the s-plane, sketch the magnitude of the frequency response of the system.

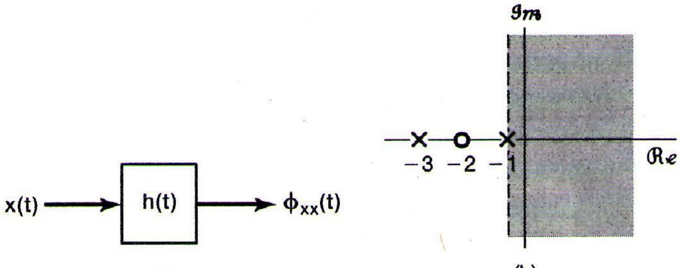
9.61. The autocorrelation function of a signal $x(t)$ is defined as

$$\phi_{xx}(\tau) = \int_{-\infty}^{\infty} x(t)x(t+\tau)\,dt.$$

(a) Determine, in terms of $x(t)$, the impulse response $h(t)$ of an LTI system for which, when the input is $x(t)$, the output is $\phi_{xx}(t)$ [Figure P9.61(a)].

(b) From your answer in part (a), determine $\Phi_{xx}(s)$, the Laplace transform of $\phi_{xx}(\tau)$ in terms of $X(s)$. Also, express $\Phi_{xx}(j\omega)$, the Fourier transform of $\phi_{xx}(\tau)$, in terms of $X(j\omega)$.

(c) If $X(s)$ has the pole-zero pattern and ROC shown in Figure P9.61(b), sketch the pole-zero pattern and indicate the ROC for $\Phi_{xx}(s)$.



(a) (b)

Figure P9.61

9.62. In a number of applications in signal design and analysis, the class of signals

$$\phi_n(t) = e^{-t/2}L_n(t)u(t), \quad n = 0, 1, 2, \ldots, \tag{P9.62–1}$$

where

$$L_n(t) = \frac{e^t}{n!}\frac{d^n}{dt^n}(t^n e^{-t}), \tag{P9.62–2}$$

is encountered.

(a) The functions $L_n(t)$ are referred to as *Laguerre polynomials*. To verify that they in fact have the form of polynomials, determine $L_0(t)$, $L_1(t)$, and $L_2(t)$ explicitly.

(b) Using the properties of the Laplace transform in Table 9.1 and Laplace transform pairs in Table 9.2, determine the Laplace transform $\Phi_n(s)$ of $\phi_n(t)$.

(c) The set of signals $\phi_n(t)$ can be generated by exciting a network of the form in Figure P9.62 with an impulse. From your result in part (b), determine $H_1(s)$ and $H_2(s)$ so that the impulse responses along the cascade chain are the signals $\phi_n(t)$ as indicated.

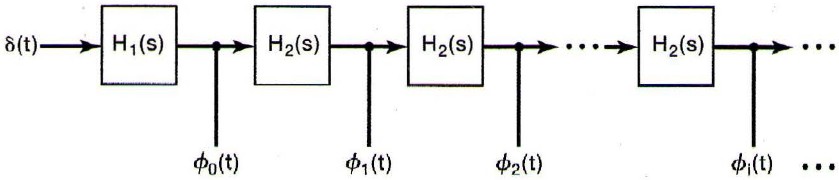


Figure P9.62

9.63. In filter design, it is often possible and convenient to transform a lowpass filter to a highpass filter and vice versa. With $H(s)$ denoting the transfer function of the original filter and $G(s)$ that of the transformed filter, one such commonly used transformation consists of replacing s by $1/s$; that is,

$$G(s) = H\left(\frac{1}{s}\right).$$

(a) For $H(s) = 1/(s + 1/2)$, sketch $|H(j\omega)|$ and $|G(j\omega)|$.

(b) Determine the linear constant-coefficient differential equation associated with $H(s)$ and with $G(s)$.

(c) Now consider a more general case in which $H(s)$ is the transfer function associated with the linear constant-coefficient differential equation in the general form

$$\sum_{k=0}^{N} a_k \frac{d^k y(t)}{dt^k} = \sum_{k=0}^{N} b_k \frac{d^k x(t)}{dt^k}. \tag{P9.63–1}$$

Without any loss of generality, we have assumed that the number of derivatives N is the same on both sides of the equation, although in any particular case, some of the coefficients may be zero. Determine $H(s)$ and $G(s)$.

(d) From your result in part (c), determine, in terms of the coefficients in eq. (P9.63–1), the linear constant-coefficient differential equation associated with $G(s)$.

9.64. Consider the *RLC* circuit shown in Figure 9.27 with input $x(t)$ and output $y(t)$.

(a) Show that if R, L, and C are all positive, then this LTI system is stable.

(b) How should R, L, and C be related to each other so that the system represents a second-order Butterworth filter?

9.65. (a) Determine the differential equation relating $v_i(t)$ and $v_0(t)$ for the *RLC* circuit of Figure P9.65.

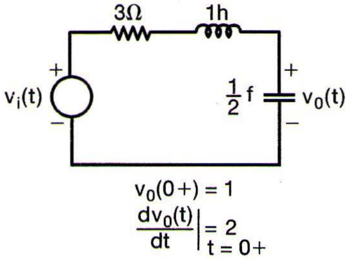


Figure P9.65

(b) Suppose that $v_i(t) = e^{-3t}u(t)$. Using the unilateral Laplace transform, determine $v_o(t)$ for $t > 0$.

9.66. Consider the *RL* circuit shown in Figure P9.66. Assume that the current $i(t)$ has reached a steady state with the switch at position *A*. At time $t = 0$, the switch is moved from position *A* to position *B*.

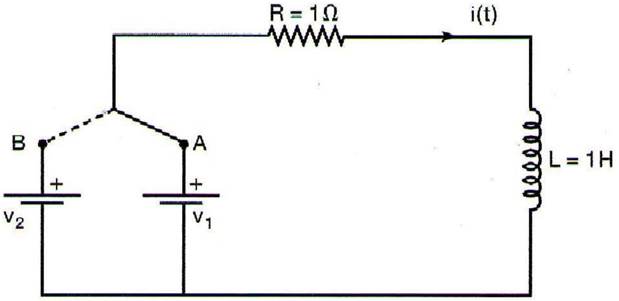


Figure P9.66

(a) Find the differential equation relating $i(t)$ and v_2 for $t > 0^-$. Specify the initial condition (i.e., the value of $i(0^-)$) for this differential equation in terms of v_1.

(b) Using the properties of the unilateral Laplace transform in Table 9.3, determine and plot the current $i(t)$ for each of the following values of v_1 and v_2:

(i) $v_1 = 0\ V, v_2 = 2\ V$

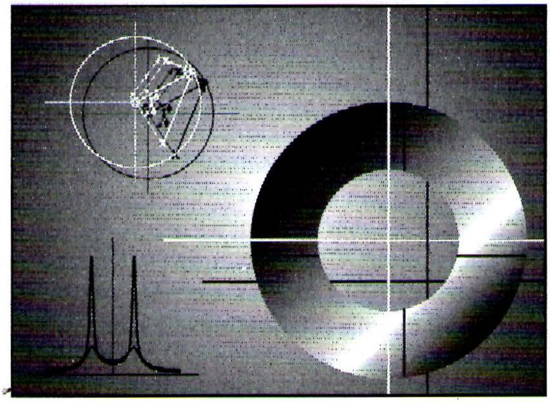
(ii) $v_1 = 4\ V, v_2 = 0\ V$

(iii) $v_1 = 4\ V, v_2 = 2\ V$

Using your answers for (i), (ii), and (iii), argue that the current $i(t)$ may be expressed as a sum of the circuit's zero-state response and zero-input response.

10

THE z-TRANSFORM



10.0 INTRODUCTION

In Chapter 9, we developed the Laplace transform as an extension of the continuous-time Fourier transform. This extension was motivated in part by the fact that it can be applied to a broader class of signals than the Fourier transform can, since there are many signals for which the Fourier transform does not converge but the Laplace transform does. The Laplace transform allowed us, for example, to perform transform analysis of unstable systems and to develop additional insights and tools for LTI system analysis.

In this chapter, we use the same approach for discrete time as we develop the z-transform, which is the discrete-time counterpart of the Laplace transform. As we will see, the motivations for and properties of the z-transform closely parallel those of the Laplace transform. Just as with the relationship between continuous-time and discrete-time Fourier transforms, however, we will encounter some important distinctions between the z-transform and the Laplace transform that arise from the fundamental differences between continuous-time and discrete-time signals and systems.

10.1 THE z-TRANSFORM

As we saw in Section 3.2, for a discrete-time linear time-invariant system with impulse response $h[n]$, the response $y[n]$ of the system to a complex exponential input of the form z^n is

$$y[n] = H(z)z^n, \tag{10.1}$$

where

$$H(z) = \sum_{n=-\infty}^{+\infty} h[n]z^{-n}. \tag{10.2}$$

For $z = e^{j\omega}$ with ω real (i.e., with $|z| = 1$), the summation in eq. (10.2) corresponds to the discrete-time Fourier transform of $h[n]$. More generally, when $|z|$ is not restricted to unity, the summation is referred to as the *z-transform* of $h[n]$.

The z-transform of a general discrete-time signal $x[n]$ is defined as[1]

$$\boxed{X(z) \triangleq \sum_{n=-\infty}^{+\infty} x[n]z^{-n},} \tag{10.3}$$

where z is a complex variable. For convenience, the z-transform of $x[n]$ will sometimes be denoted as $\mathcal{Z}\{x[n]\}$ and the relationship between $x[n]$ and its z-transform indicated as

$$x[n] \overset{\mathcal{Z}}{\longleftrightarrow} X(z). \tag{10.4}$$

In Chapter 9, we considered a number of important relationships between the Laplace transform and the Fourier transform for continuous-time signals. In a similar, but not identical, way, there are a number of important relationships between the z-transform and the discrete-time Fourier transform. To explore these relationships, we express the complex variable z in polar form as

$$z = re^{j\omega}, \tag{10.5}$$

with r as the magnitude of z and ω as the angle of z. In terms of r and ω, eq. (10.3) becomes

$$X(re^{j\omega}) = \sum_{n=-\infty}^{+\infty} x[n](re^{j\omega})^{-n},$$

or equivalently,

$$X(re^{j\omega}) = \sum_{n=-\infty}^{+\infty} \{x[n]r^{-n}\}e^{-j\omega n}. \tag{10.6}$$

From eq. (10.6), we see that $X(re^{j\omega})$ is the Fourier transform of the sequence $x[n]$ multiplied by a real exponential r^{-n}; that is,

$$X(re^{j\omega}) = \mathfrak{F}\{x[n]r^{-n}\}. \tag{10.7}$$

The exponential weighting r^{-n} may be decaying or growing with increasing n, depending on whether r is greater than or less than unity. We note in particular that, for $r = 1$, or equivalently, $|z| = 1$, eq. (10.3) reduces to the Fourier transform; that is,

---

[1] The z-transform defined in eq. (10.3) is often referred to as the *bilateral* z-transform, to distinguish it from the *unilateral* z-transform, which we develop in Section 10.9. The bilateral z-transform involves a summation from $-\infty$ to $+\infty$, while the unilateral transform has a form similar to eq. (10.3), but with summation limits from 0 to $+\infty$. Since we are mostly concerned with the bilateral z-transform, we will refer to $X(z)$ as defined in eq. (10.3) simply as the z-transform, except in Section 10.9, in which we use the words "unilateral" and "bilateral" to avoid ambiguity.

$$X(z)\Big|_{z=e^{j\omega}} = X(e^{j\omega}) = \mathfrak{F}\{x[n]\}. \tag{10.8}$$

The relationship between the z-transform and Fourier transform for discrete-time signals parallels closely the corresponding discussion in Section 9.1 for continuous-time signals, but with some important differences. In the continuous-time case, the Laplace transform reduces to the Fourier transform when the real part of the transform variable is zero. Interpreted in terms of the s-plane, this means that the Laplace transform reduces to the Fourier transform on the imaginary axis (i.e., for $s = j\omega$). In contrast, the z-transform reduces to the Fourier transform when the magnitude of the transform variable z is *unity* (i.e, for $z = e^{j\omega}$). Thus, the z-transform reduces to the Fourier transform on the contour in the complex z-plane corresponding to a circle with a radius of unity, as indicated in Figure 10.1. This circle in the z-plane is referred to as the *unit circle* and plays a role in the discussion of the z-transform similar to the role of the imaginary axis in the s-plane for the Laplace transform.

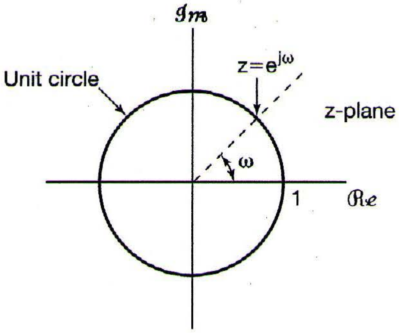


Figure 10.1 Complex z-plane. The z-transform reduces to the Fourier transform for values of z on the unit circle.

From eq. (10.7), we see that, for convergence of the z-transform, we require that the Fourier transform of $x[n]r^{-n}$ converge. For any specific sequence $x[n]$, we would expect this convergence for some values of r and not for others. In general, the z-transform of a sequence has associated with it a range of values of z for which $X(z)$ converges. As with the Laplace transform, this range of values is referred to as the *region of convergence* (ROC). If the ROC includes the unit circle, then the Fourier transform also converges. To illustrate the z-transform and the associated region of convergence, let us consider several examples.

Example 10.1

Consider the signal $x[n] = a^n u[n]$. Then, from eq. (10.3),

$$X(z) = \sum_{n=-\infty}^{+\infty} a^n u[n] z^{-n} = \sum_{n=0}^{\infty} (az^{-1})^n.$$

For convergence of $X(z)$, we require that $\sum_{n=0}^{\infty} |az^{-1}|^n < \infty$. Consequently, the region of convergence is the range of values of z for which $|az^{-1}| < 1$, or equivalently, $|z| > |a|$.

Then

$$X(z) = \sum_{n=0}^{\infty}(az^{-1})^n = \frac{1}{1-az^{-1}} = \frac{z}{z-a}, \quad |z| > |a|. \tag{10.9}$$

Thus, the z-transform for this signal is well-defined for any value of a, with an ROC determined by the magnitude of a according to eq. (10.9). For example, for $a = 1$, $x[n]$ is the unit step sequence with z-transform

$$X(z) = \frac{1}{1-z^{-1}}, \quad |z| > 1.$$

We see that the z-transform in eq. (10.9) is a rational function. Consequently, just as with rational Laplace transforms, the z-transform can be characterized by its zeros (the roots of the numerator polynomial) and its poles (the roots of the denominator polynomial). For this example, there is one zero, at $z = 0$, and one pole, at $z = a$. The pole-zero plot and the region of convergence for Example 10.1 are shown in Figure 10.2 for a value of a between 0 and 1. For $|a| > 1$, the ROC does not include the unit circle, consistent with the fact that, for these values of a, the Fourier transform of $a^n u[n]$ does not converge.

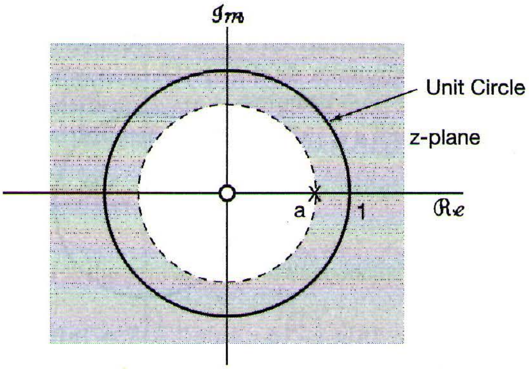


Figure 10.2 Pole-zero plot and region of convergence for Example 10.1 for $0 < a < 1$.

Example 10.2

Now let $x[n] = -a^n u[-n-1]$. Then

$$\begin{aligned} X(z) &= -\sum_{n=-\infty}^{+\infty} a^n u[-n-1] z^{-n} = -\sum_{n=-\infty}^{-1} a^n z^{-n} \\ &= -\sum_{n=1}^{\infty} a^{-n} z^n = 1 - \sum_{n=0}^{\infty} (a^{-1}z)^n. \end{aligned} \tag{10.10}$$

If $|a^{-1}z| < 1$, or equivalently, $|z| < |a|$, the sum in eq. (10.10) converges and

$$X(z) = 1 - \frac{1}{1-a^{-1}z} = \frac{1}{1-az^{-1}} = \frac{z}{z-a}, \quad |z| < |a|. \tag{10.11}$$

The pole-zero plot and region of convergence for this example are shown in Figure 10.3 for a value of a between 0 and 1.

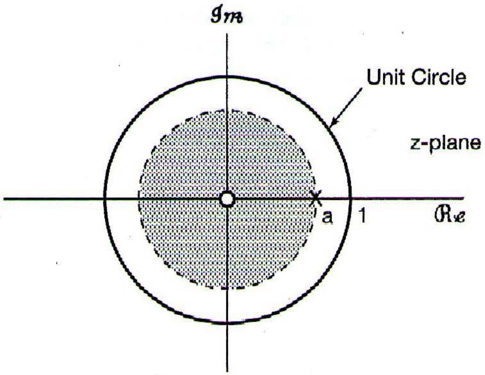


Figure 10.3 Pole-zero plot and region of convergence for Example 10.2 for $0 < a < 1$.

Comparing eqs. (10.9) and (10.11), and Figures 10.2 and 10.3, we see that the algebraic expression for $X(z)$ and the corresponding pole-zero plot are identical in Examples 10.1 and 10.2, and the z-transforms differ only in their regions of convergence. Thus, as with the Laplace transform, specification of the z-transform requires both the algebraic expression and the region of convergence. Also, in both examples, the sequences were exponentials and the resulting z-transforms were rational. In fact, as further suggested by the following examples, $X(z)$ will be rational whenever $x[n]$ is a linear combination of real or complex exponentials:

Example 10.3

Let us consider a signal that is the sum of two real exponentials:

$$x[n] = 7\left(\frac{1}{3}\right)^n u[n] - 6\left(\frac{1}{2}\right)^n u[n]. \tag{10.12}$$

The z-transform is then

$$\begin{aligned} X(z) &= \sum_{n=-\infty}^{+\infty}\left\{7\left(\frac{1}{3}\right)^n u[n] - 6\left(\frac{1}{2}\right)^n u[n]\right\} z^{-n} \\ &= 7\sum_{n=-\infty}^{+\infty}\left(\frac{1}{3}\right)^n u[n]z^{-n} - 6\sum_{n=-\infty}^{+\infty}\left(\frac{1}{2}\right)^n u[n]z^{-n} \\ &= 7\sum_{n=0}^{\infty}\left(\frac{1}{3}z^{-1}\right)^n - 6\sum_{n=0}^{+\infty}\left(\frac{1}{2}z^{-1}\right)^n \qquad (10.13) \\ &= \frac{7}{1-\frac{1}{3}z^{-1}} - \frac{6}{1-\frac{1}{2}z^{-1}} = \frac{1-\frac{3}{2}z^{-1}}{(1-\frac{1}{3}z^{-1})(1-\frac{1}{2}z^{-1})} \qquad (10.14) \\ &= \frac{z(z-\frac{3}{2})}{(z-\frac{1}{3})(z-\frac{1}{2})}. \qquad (10.15) \end{aligned}$$

For convergence of $X(z)$, both sums in eq. (10.13) must converge, which requires that both $|(1/3)z^{-1}| < 1$ and $|(1/2)z^{-1}| < 1$, or equivalently, $|z| > 1/3$ and $|z| > 1/2$. Thus, the region of convergence is $|z| > 1/2$.

The z-transform for this example can also be obtained using the results of Example 10.1. Specifically, from the definition of the z-transform in eq. (10.3), we see that the z-transform is linear; that is, if $x[n]$ is the sum of two terms, then $X(z)$ will be the sum of the z-transforms of the individual terms and will converge when both z-transforms converge. From Example 10.1,

$$\left(\frac{1}{3}\right)^n u[n] \overset{z}{\longleftrightarrow} \frac{1}{1-\frac{1}{3}z^{-1}}, \quad |z| > \frac{1}{3} \tag{10.16}$$

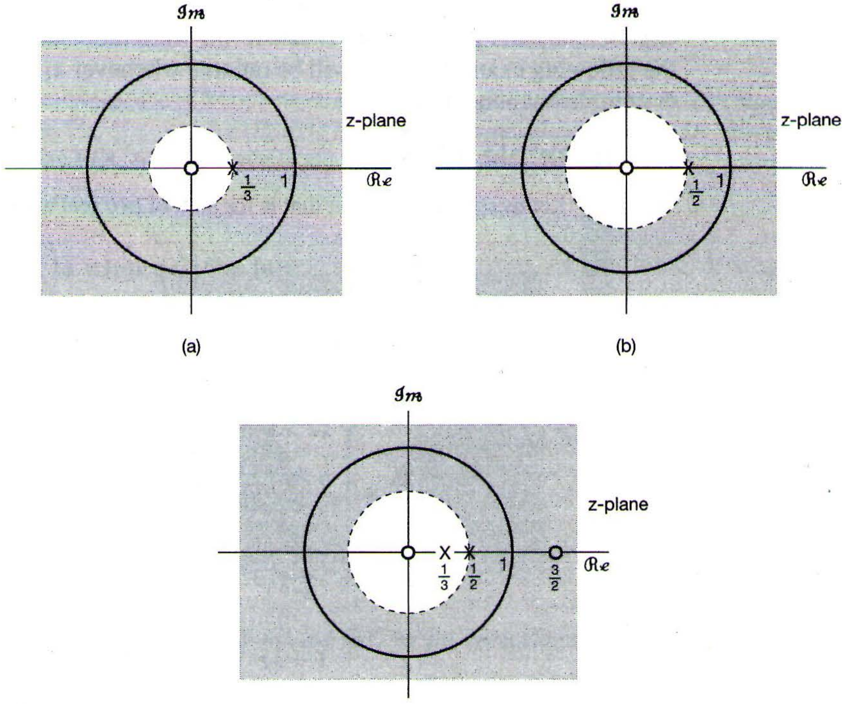
and

$$\left(\frac{1}{2}\right)^n u[n] \overset{z}{\longleftrightarrow} \frac{1}{1-\frac{1}{2}z^{-1}}, \quad |z| > \frac{1}{2}, \tag{10.17}$$

and consequently,

$$7\left(\frac{1}{3}\right)^n u[n] - 6\left(\frac{1}{2}\right)^n u[n] \overset{z}{\longleftrightarrow} \frac{7}{1-\frac{1}{3}z^{-1}} - \frac{6}{1-\frac{1}{2}z^{-1}}, \quad |z| > \frac{1}{2}, \tag{10.18}$$

as we determined before. The pole-zero plot and ROC for the z-transform of each of the individual terms and for the combined signal are shown in Figure 10.4.



(a) (b)

(c)

Figure 10.4 Pole-zero plot and region of convergence for the individual terms and the sum in Example 10.3: (a) $1/(1-\frac{1}{3}z^{-1})$, $|z| > \frac{1}{3}$; (b) $1/(1-\frac{1}{2}z^{-1})$, $|z| > \frac{1}{2}$; (c) $7/(1-\frac{1}{3}z^{-1}) - 6/(1-\frac{1}{2}z^{-1})$, $|z| > \frac{1}{2}$.

Example 10.4

Let us consider the signal

$$\begin{aligned} x[n] &= (\tfrac{1}{3})^n \sin(\tfrac{\pi}{4}n)u[n] \\ &= \tfrac{1}{2j}(\tfrac{1}{3}e^{j\pi/4})^n u[n] - \tfrac{1}{2j}(\tfrac{1}{3}e^{-j\pi/4})^n u[n]. \end{aligned}$$

The z-transform of this signal is

$$\begin{aligned} X(z) &= \sum_{n=-\infty}^{\infty} \left\{ \frac{1}{2j}\left(\frac{1}{3}e^{j\pi/4}\right)^n u[n] - \frac{1}{2j}\left(\frac{1}{3}e^{-j\pi/4}\right)^n u[n] \right\} z^{-n} \\ &= \frac{1}{2j}\sum_{n=0}^{\infty}\left(\frac{1}{3}e^{j\pi/4}z^{-1}\right)^n - \frac{1}{2j}\sum_{n=0}^{\infty}\left(\frac{1}{3}e^{-j\pi/4}z^{-1}\right)^n \qquad (10.19) \\ &= \frac{1}{2j}\frac{1}{1-\frac{1}{3}e^{j\pi/4}z^{-1}} - \frac{1}{2j}\frac{1}{1-\frac{1}{3}e^{-j\pi/4}z^{-1}}, \end{aligned}$$

or equivalently,

$$X(z) = \frac{\frac{1}{3\sqrt{2}}z}{(z-\frac{1}{3}e^{j\pi/4})(z-\frac{1}{3}e^{-j\pi/4})} \qquad (10.20)$$

For convergence of $X(z)$, both sums in eq. (10.19) must converge, which requires that $|(1/3)e^{j\pi/4}z^{-1}| < 1$ and $|(1/3)e^{-j\pi/4}z^{-1}| < 1$, or equivalently, $|z| > 1/3$. The pole-zero plot and ROC for this example are shown in Figure 10.5.

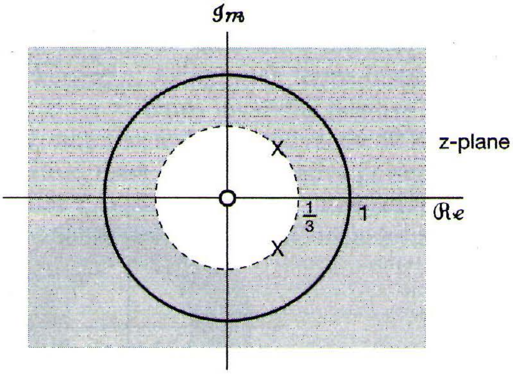


Figure 10.5 Pole-zero plot and ROC for the *z*-transform in Example 10.4.

In each of the preceding four examples, we expressed the z-transform both as a ratio of polynomials in z and as a ratio of polynomials in z^{-1}. From the definition of the z-transform as given in eq. (10.3), we see that, for sequences which are zero for $n < 0$, $X(z)$ involves only negative powers of z. Thus, for this class of signals, it is particularly convenient for $X(z)$ to be expressed in terms of polynomials in z^{-1} rather than z, and when appropriate, we will use that form in our discussion. However, reference to the poles and zeros is always in terms of the roots of the numerator and denominator expressed as polynomials in z. Also, it is sometimes convenient to refer to $X(z)$, written as a ratio of polynomials in z, as having poles at infinity if the degree of the numerator exceeds the degree of the denominator or zeros at infinity if the numerator is of smaller degree than the denominator.

10.2 THE REGION OF CONVERGENCE FOR THE z-TRANSFORM

In Chapter 9, we saw that there were specific properties of the region of convergence of the Laplace transform for different classes of signals and that understanding these properties led to further insights about the transform. In a similar manner, we explore a number of properties of the region of convergence for the z-transform. Each of the following properties and its justification closely parallel the corresponding property in Section 9.2.

Property 1: The ROC of $X(z)$ consists of a ring in the z-plane centered about the origin.

This property is illustrated in Figure 10.6 and follows from the fact that the ROC consists of those values of $z = re^{j\omega}$ for which $x[n]r^{-n}$ has a Fourier transform that converges. That is, the ROC of the z-transform of $x[n]$ consists of the values of z for which $x[n]r^{-n}$ is absolutely summable:[2]

$$\sum_{n=-\infty}^{+\infty} |x[n]|r^{-n} < \infty. \tag{10.21}$$

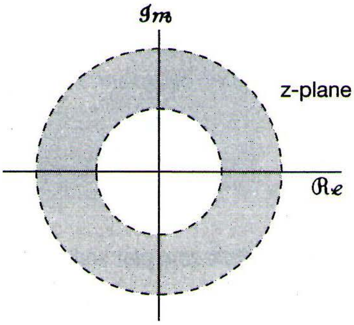


Figure 10.6 ROC as a ring in the z-plane. In some cases, the inner boundary can extend inward to the origin, in which case the ROC becomes a disc. In other cases, the outer boundary can extend outward to infinity.

---

[2] For a thorough treatment of the mathematical properties of z-transforms, see R.V. Churchill and J.W. Brown, *Complex Variables and Applications* (5th ed.) (New York: McGraw–Hill, 1990), and E. I. Jury, *Theory and Application of the z-Transform Method* (Malabar, FL: R. E. Krieger Pub. Co., 1982).

Thus, convergence is dependent only on $r = |z|$ and not on ω. Consequently, if a specific value of z is in the ROC, then all values of z on the same circle (i.e., with the same magnitude) will be in the ROC. This by itself guarantees that the ROC will consist of concentric rings. As we will see when we discuss Property 6, the ROC must in fact consist of only a single ring. In some cases the inner boundary of the ROC may extend inward to the origin, and in some cases the outer boundary may extend outward to infinity.

Property 2: The ROC does not contain any poles.

As with the Laplace transform, this property is simply a consequence of the fact that at a pole $X(z)$ is infinite and therefore, by definition, does not converge.

Property 3: If $x[n]$ is of finite duration, then the ROC is the entire z-plane, except possibly $z = 0$ and/or $z = \infty$.

A finite-duration sequence has only a finite number of nonzero values, extending, say, from $n = N_1$ to $n = N_2$, where N_1 and N_2 are finite. Thus, the z-transform is the sum of a finite number of terms; that is,

$$X(z) = \sum_{n=N_1}^{N_2} x[n]z^{-n}. \tag{10.22}$$

For z not equal to zero or infinity, each term in the sum will be finite, and consequently, $X(z)$ will converge. If N_1 is negative and N_2 positive, so that $x[n]$ has nonzero values both for $n < 0$ and $n > 0$, then the summation includes terms with both positive powers of z and negative powers of z. As $|z| \rightarrow 0$, terms involving negative powers of z become unbounded, and as $|z| \rightarrow \infty$, terms involving positive powers of z become unbounded. Consequently, for N_1 negative and N_2 positive, the ROC does not include $z = 0$ or $z = \infty$. If N_1 is zero or positive, there are only negative powers of z in eq. (10.22), and consequently, the ROC includes $z = \infty$. If N_2 is zero or negative, there are only positive powers of z in eq. (10.22), and consequently, the ROC includes $z = 0$.

Example 10.5

Consider the unit impulse signal $\delta[n]$. Its z-transform is given by

$$\delta[n] \overset{z}{\longleftrightarrow} \sum_{n=-\infty}^{+\infty} \delta[n]z^{-n} = 1, \tag{10.23}$$

with an ROC consisting of the entire z-plane, including $z = 0$ and $z = \infty$. On the other hand, consider the delayed unit impulse $\delta[n-1]$, for which

$$\delta[n-1] \overset{z}{\longleftrightarrow} \sum_{n=-\infty}^{+\infty} \delta[n-1]z^{-n} = z^{-1}. \tag{10.24}$$

This z-transform is well defined except at $z = 0$, where there is a pole. Thus, the ROC consists of the entire z-plane, including $z = \infty$ but excluding $z = 0$. Similarly, consider an impulse advanced in time, namely, $\delta[n+1]$. In this case,

$$\delta[n+1] \overset{z}{\longleftrightarrow} \sum_{n=-\infty}^{+\infty} \delta[n+1]z^{-n} = z, \tag{10.25}$$

which is well defined for all finite values of z. Thus, the ROC consists of the entire finite z-plane (including $z = 0$), but there is a pole at infinity.

Property 4: If $x[n]$ is a right-sided sequence, and if the circle $|z| = r_0$ is in the ROC, then all finite values of z for which $|z| > r_0$ will also be in the ROC.

The justification for this property follows in a manner identical to that of Property 4 in Section 9.2. A right-sided sequence is zero prior to some value of n, say, N_1. If the circle $|z| = r_0$ is in the ROC, then $x[n]r_0^{-n}$ is absolutely summable. Now consider $|z| = r_1$ with $r_1 > r_0$, so that r_1^{-n} decays more quickly than r_0^{-n} for increasing n. As illustrated in Figure 10.7, this more rapid exponential decay will further attenuate sequence values

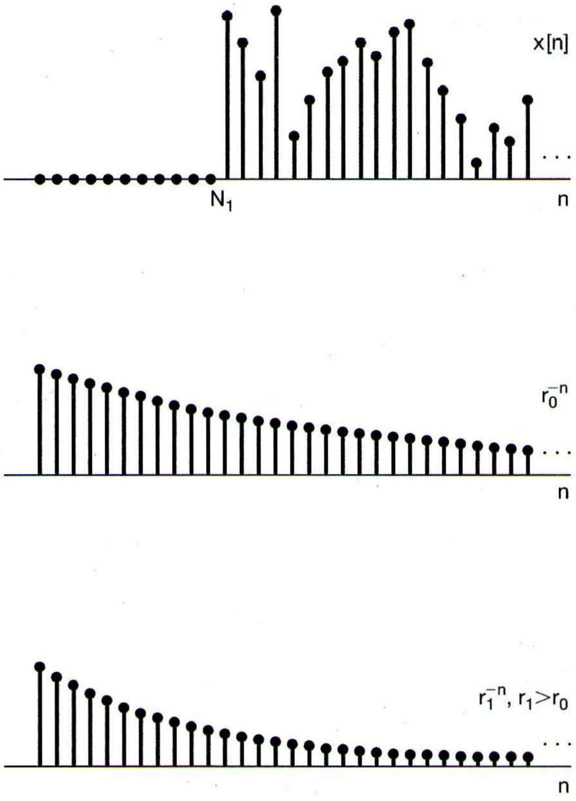


Figure 10.7 With $r_1 > r_0, x[n]r_1^{-n}$ decays faster with increasing n than does $x[n]r_0^{-n}$. Since $x[n] = 0, n < N_1$, this implies that if $x[n]r_0^{-n}$ is absolutely summable, then $x[n]r_1^{-n}$ will be also.

for positive values of n and cannot cause sequence values for negative values of n to become unbounded, since $x[n]$ is right sided and, in particular, $x[n]z^{-n} = 0$ for $n < N_1$. Consequently, $x[n]r_1^{-n}$ is absolutely summable.

For right-sided sequences in general, eq. (10.3) takes the form

$$X(z) = \sum_{n=N_1}^{\infty} x[n]z^{-n}, \tag{10.26}$$

where N_1 is finite and may be positive or negative. If N_1 is negative, then the summation in eq. (10.26) includes terms with positive powers of z, which become unbounded as $|z| \to \infty$. Consequently, for right-sided sequences in general, the ROC will not include infinity. However, for the particular class of causal sequences, i.e., sequences that are zero for $n < 0$, N_1 will be nonnegative, and consequently, the ROC will include $z = \infty$.

Property 5: If $x[n]$ is a left-sided sequence, and if the circle $|z| = r_0$ is in the ROC, then all values of z for which $0 < |z| < r_0$ will also be in the ROC.

Again, this property closely parallels the corresponding property for Laplace transforms, and the proof of it and its basis in intuition are similar to the proof and intuition for Property 4. In general, for left-sided sequences, from eq. (10.3), the summation for the z-transform will be of the form

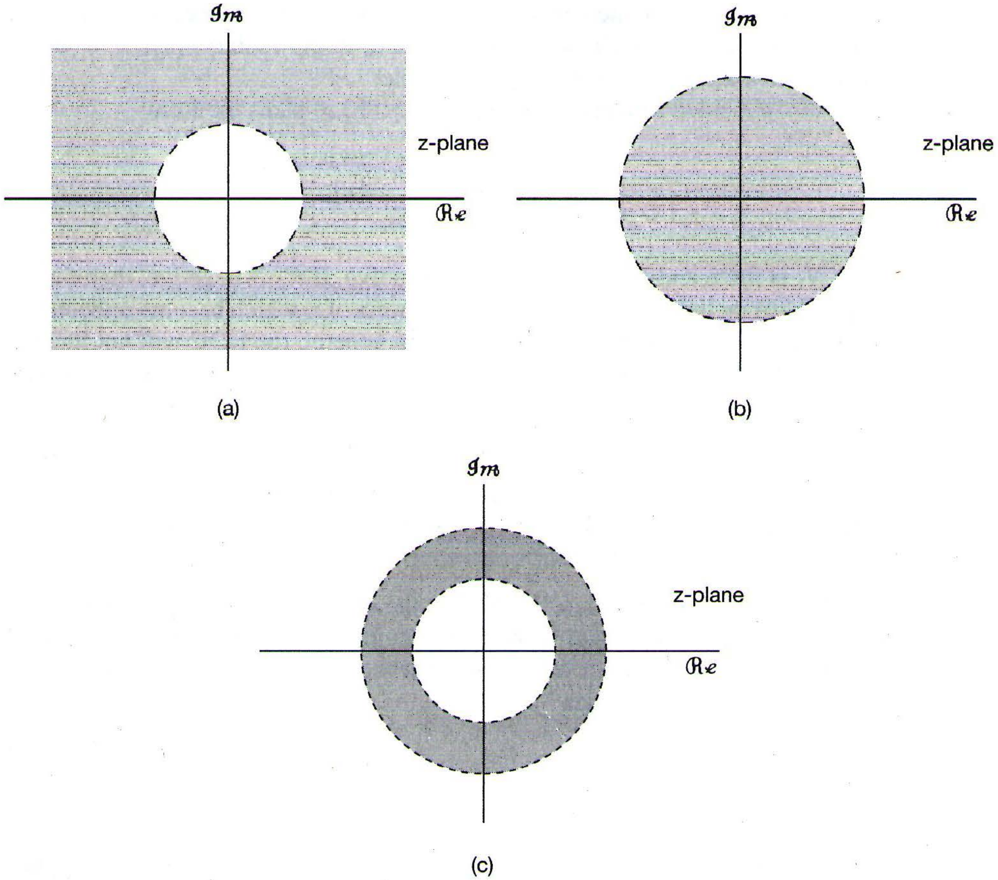
$$X(z) = \sum_{n=-\infty}^{N_2} x[n]z^{-n}, \tag{10.27}$$

where N_2 may be positive or negative. If N_2 is positive, then eq. (10.27) includes negative powers of z, which become unbounded as $|z| \to 0$. Consequently, for left-sided sequences, the ROC will not in general include $z = 0$. However, if $N_2 \leq 0$ (so that $x[n] = 0$ for all $n > 0$), the ROC will include $z = 0$.

Property 6: If $x[n]$ is two sided, and if the circle $|z| = r_0$ is in the ROC, then the ROC will consist of a ring in the z-plane that includes the circle $|z| = r_0$.

As with Property 6 in Section 9.2, the ROC for a two-sided signal can be examined by expressing $x[n]$ as the sum of a right-sided and a left-sided signal. The ROC for the right-sided component is a region bounded on the inside by a circle and extending outward to (and possibly including) infinity. The ROC for the left-sided component is a region bounded on the outside by a circle and extending inward to, and possibly including, the origin. The ROC for the composite signal includes the intersection of the ROCs of the components. As illustrated in Figure 10.8, the overlap (assuming that there is one) is a ring in the z-plane.

We illustrate the foregoing properties with examples that closely parallel Examples 9.6 and 9.7.


$\mathcal{Im}$
z-plane
$\mathcal{Re}$
$\mathcal{Im}$
z-plane
$\mathcal{Re}$
$\mathcal{Im}$
z-plane
$\mathcal{Re}$

(a) (b) (c)

Figure 10.8 (a) ROC for right-sided sequence; (b) ROC for left-sided sequence; (c) intersection of the ROCs in (a) and (b), representing the ROC for a two-sided sequence that is the sum of the right-sided and the left-sided sequence.

Example 10.6

Consider the signal

$$x[n] = \begin{cases} a^n, & 0 \le n \le N-1, \;\; a > 0 \\ 0, & \text{otherwise} \end{cases}.$$

Then

$$\begin{aligned} X(z) &= \sum_{n=0}^{N-1} a^n z^{-n} \\ &= \sum_{n=0}^{N-1} (az^{-1})^n \\ &= \frac{1-(az^{-1})^N}{1-az^{-1}} = \frac{1}{z^{N-1}} \frac{z^N - a^N}{z-a}. \end{aligned} \tag{10.28}$$

Since $x[n]$ is of finite length, it follows from Property 3 that the ROC includes the entire z-plane except possibly the origin and/or infinity. In fact, from our discussion of Property 3, since $x[n]$ is zero for $n < 0$, the ROC will extend to infinity. However, since $x[n]$ is nonzero for some positive values of n, the ROC will not include the origin. This is evident from eq. (10.28), from which we see that there is a pole of order $N - 1$ at $z = 0$. The N roots of the numerator polynomial are at

$$z_k = ae^{j(2\pi k/N)}, \quad k = 0, 1, \ldots, N-1. \tag{10.29}$$

The root for $k = 0$ cancels the pole at $z = a$. Consequently, there are no poles other than at the origin. The remaining zeros are at

$$z_k = ae^{j(2\pi k/N)}, \quad k = 1, \ldots, N-1. \tag{10.30}$$

The pole-zero pattern is shown in Figure 10.9.

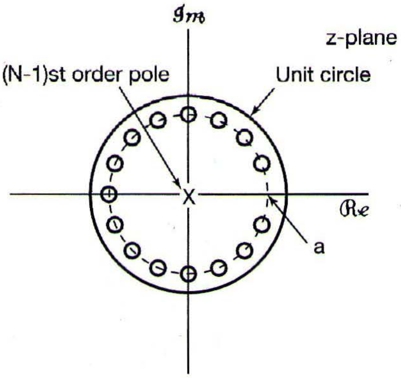


Figure 10.9 Pole-zero pattern for Example 10.6 with $N = 16$ and $0 < a < 1$. The region of convergence for this example consists of all values of z except $z = 0$.

Example 10.7

Let

$$x[n] = b^{|n|}, \quad b > 0. \tag{10.31}$$

This two-sided sequence is illustrated in Figure 10.10, for both $b < 1$ and $b > 1$. The z-transform for the sequence can be obtained by expressing it as the sum of a right-sided and a left-sided sequence. We have

$$x[n] = b^n u[n] + b^{-n} u[-n-1]. \tag{10.32}$$

From Example 10.1,

$$b^n u[n] \overset{z}{\longleftrightarrow} \frac{1}{1 - bz^{-1}}, \quad |z| > b, \tag{10.33}$$

and from Example 10.2,

$$b^{-n} u[-n-1] \overset{z}{\longleftrightarrow} \frac{-1}{1 - b^{-1}z^{-1}}, \quad |z| < \frac{1}{b}. \tag{10.34}$$

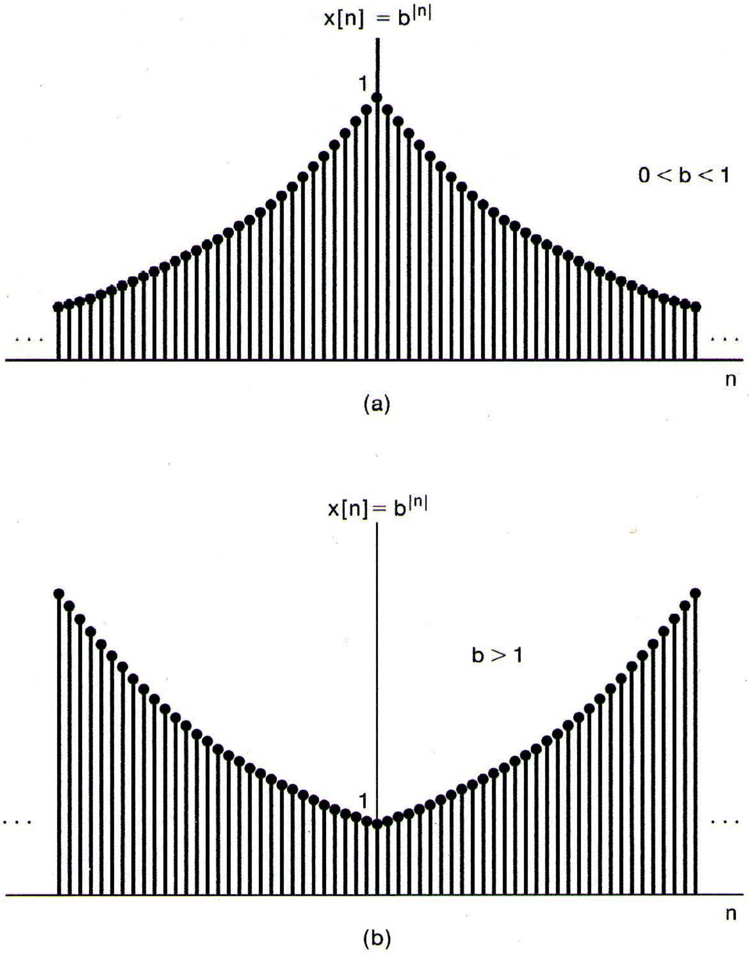


Figure 10.10 Sequence $x[n] = b^{|n|}$ for $0 < b < 1$ and for $b > 1$: (a) $b = 0.95$; (b) $b = 1.05$.

In Figures 10.11(a)–(d) we show the pole-zero pattern and ROC for eqs. (10.33) and (10.34), for values of $b > 1$ and $0 < b < 1$. For $b > 1$, there is no common ROC, and thus the sequence in eq. (10.31) will not have a z-transform, even though the right-sided and left-sided components do individually. For $b < 1$, the ROCs in eqs. (10.33) and (10.34) overlap, and thus the z-transform for the composite sequence is

$$X(z) = \frac{1}{1 - bz^{-1}} - \frac{1}{1 - b^{-1}z^{-1}}, \quad b < |z| < \frac{1}{b}, \tag{10.35}$$

or equivalently,

$$X(z) = \frac{b^2 - 1}{b} \frac{z}{(z - b)(z - b^{-1})}, \quad b < |z| < \frac{1}{b}. \tag{10.36}$$

The corresponding pole-zero pattern and ROC are shown in Figure 10.11(e).

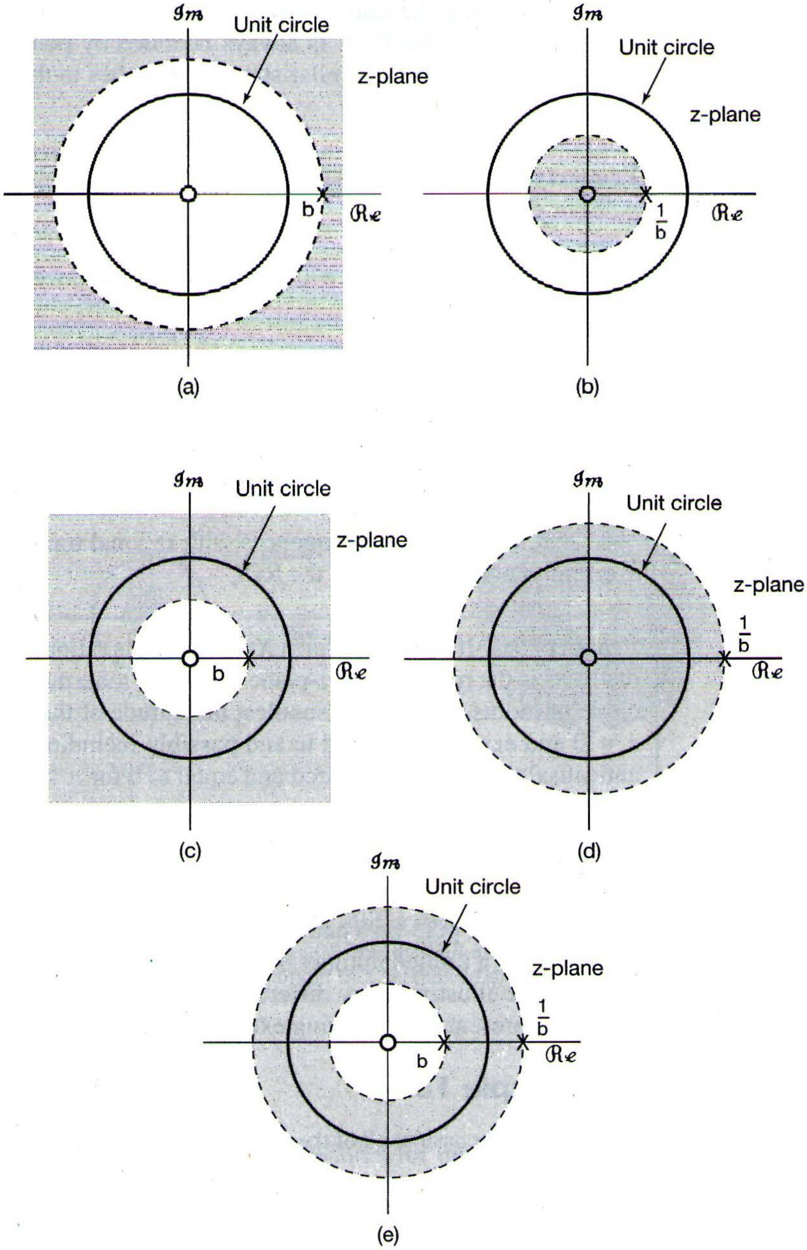

Im
Unit circle
z-plane
b
Re
(a)
Im
Unit circle
z-plane
1/b
Re
(b)
Im
Unit circle
z-plane
b
Re
(c)
Im
Unit circle
z-plane
1/b
Re
(d)
Im
Unit circle
z-plane
b
1/b
Re
(e)

Figure 10.11 Pole-zero plots and ROCs for Example 10.7: (a) eq. (10.33) for $b > 1$; (b) eq. (10.34) for $b > 1$; (c) eq. (10.33) for $0 < b < 1$; (d) eq. (10.34) for $0 < b < 1$; (e) pole-zero plot and ROC for eq. (10.36) with $0 < b < 1$. For $b > 1$, the z-transform of $x[n]$ in eq. (10.31) does not converge for any value of z.

In discussing the Laplace transform in Chapter 9, we remarked that for a rational Laplace transform, the ROC is always bounded by poles or infinity. We observe that in the foregoing examples a similar statement applies to the z-transform, and in fact, this is always true:

> **Property 7:** If the z-transform $X(z)$ of $x[n]$ is rational, then its ROC is bounded by poles or extends to infinity.

Combining Property 7 with Properties 4 and 5, we have

> **Property 8:** If the z-transform $X(z)$ of $x[n]$ is rational, and if $x[n]$ is right sided, then the ROC is the region in the z-plane outside the outermost pole—i.e., outside the circle of radius equal to the largest magnitude of the poles of $X(z)$. Furthermore, if $x[n]$ is causal (i.e., if it is right sided and equal to 0 for $n < 0$), then the ROC also includes $z = \infty$.

Thus, for right-sided sequences with rational transforms, the poles are all closer to the origin than is any point in the ROC.

> **Property 9:** If the z-transform $X(z)$ of $x[n]$ is rational, and if $x[n]$ is left sided, then the ROC is the region in the z-plane inside the innermost nonzero pole—i.e., inside the circle of radius equal to the smallest magnitude of the poles of $X(z)$ other than any at $z = 0$ and extending inward to and possibly including $z = 0$. In particular, if $x[n]$ is anticausal (i.e., if it is left sided and equal to 0 for $n > 0$), then the ROC also includes $z = 0$.

Thus, for left-sided sequences, the poles of $X(z)$ other than any at $z = 0$ are farther from the origin than is any point in the ROC.

For a given pole-zero pattern, or equivalently, a given rational algebraic expression $X(z)$, there are a limited number of different ROCs that are consistent with the preceding properties. To illustrate how different ROCs can be associated with the same pole-zero pattern, we present the following example, which closely parallels Example 9.8.

Example 10.8

Let us consider all of the possible ROCs that can be connected with the function

$$X(z) = \frac{1}{(1 - \frac{1}{3}z^{-1})(1 - 2z^{-1})}. \tag{10.37}$$

The associated pole-zero pattern is shown in Figure 10.12(a). Based on our discussion in this section, there are three possible ROCs that can be associated with this algebraic expression for the z-transform. These ROCs are indicated in Figure 10.12(b)–(d). Each corresponds to a different sequence. Figure 10.12(b) is associated with a right-sided sequence, Figure 10.12(c) with a left-sided sequence, and Figure 10.12(d) with a two-sided sequence. Since Figure 10.12(d) is the only one for which the ROC includes the unit circle, the sequence corresponding to this choice of ROC is the only one of the three for which the Fourier transform converges.

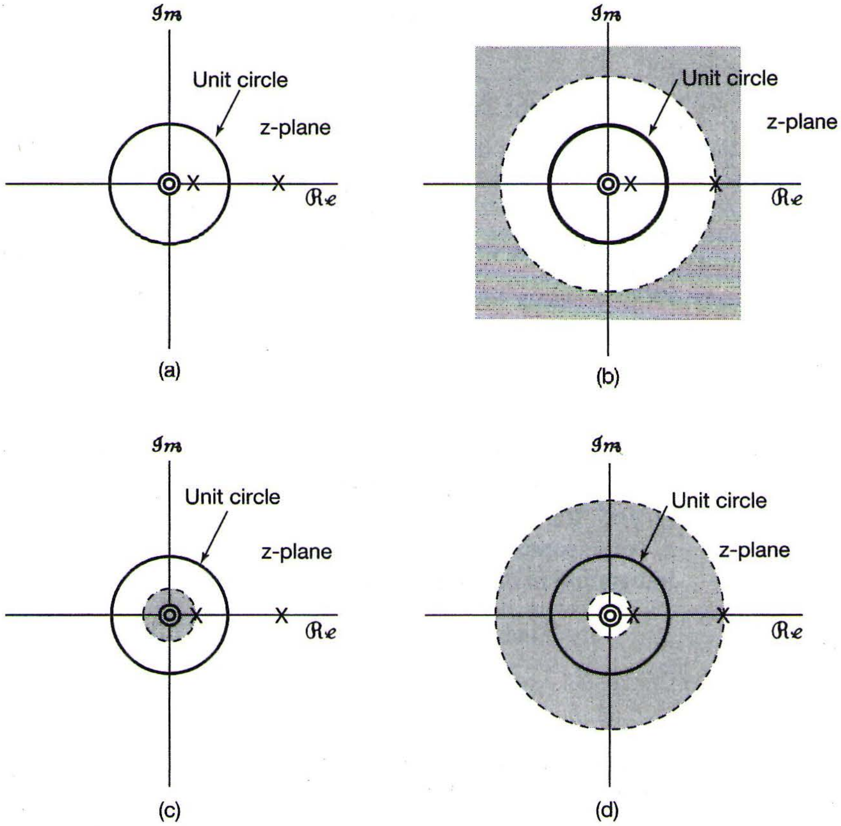

$\mathcal{Im}$
Unit circle
z-plane
$\mathcal{Re}$
(a)
$\mathcal{Im}$
Unit circle
z-plane
$\mathcal{Re}$
(b)
$\mathcal{Im}$
Unit circle
z-plane
$\mathcal{Re}$
(c)
$\mathcal{Im}$
Unit circle
z-plane
$\mathcal{Re}$
(d)

Figure 10.12 The three possible ROCs that can be connected with the expression for the z-transform in Example 10.8: (a) pole-zero pattern for $X(z)$; (b) pole-zero pattern and ROC if $x[n]$ is right sided; (c) pole-zero pattern and ROC if $x[n]$ is left sided; (d) pole-zero pattern and ROC if $x[n]$ is two sided. In each case, the zero at the origin is a second-order zero.

10.3 THE INVERSE z-TRANSFORM

In this section, we consider several procedures for determining a sequence when its z-transform is known. To begin, let us consider the formal relation expressing a sequence in terms of its z-transform. This expression can be obtained on the basis of the interpretation, developed in Section 10.1, of the z-transform as the Fourier transform of an exponentially weighted sequence. Specifically, as expressed in eq. (10.7),

$$X(re^{j\omega}) = \mathfrak{F}\{x[n]r^{-n}\}, \tag{10.38}$$

for any value of r so that $z = re^{j\omega}$ is inside the ROC. Applying the inverse Fourier transform to both sides of eq. (10.38) yields

$$x[n]r^{-n} = \mathfrak{F}^{-1}\{X(re^{j\omega})\},$$

or

$$x[n] = r^n \mathfrak{F}^{-1}[X(re^{j\omega})]. \tag{10.39}$$

Using the inverse Fourier transform expression in eq. (5.8), we have

$$x[n] = r^n \frac{1}{2\pi} \int_{2\pi} X(re^{j\omega}) e^{j\omega n} d\omega,$$

or, moving the exponential factor r^n inside the integral and combining it with the term $e^{j\omega n}$,

$$x[n] = \frac{1}{2\pi} \int_{2\pi} X(re^{j\omega})(re^{j\omega})^n d\omega. \tag{10.40}$$

That is, we can recover $x[n]$ from its z-transform evaluated along a contour $z = re^{j\omega}$ in the ROC, with r fixed and ω varying over a 2π interval. Let us now change the variable of integration from ω to z. With $z = re^{j\omega}$ and r fixed, $dz = jre^{j\omega} d\omega = jz d\omega$, or $d\omega = (1/j)z^{-1} dz$. The integration in eq. (10.40) is over a 2π interval in ω, which, in terms of z, corresponds to one traversal around the circle $|z| = r$. Consequently, in terms of an integration in the z-plane, eq. (10.40) can be rewritten as

$$\boxed{x[n] = \frac{1}{2\pi j} \oint X(z) z^{n-1} dz,} \tag{10.41}$$

where the symbol $\bigcirc$ denotes integration around a counterclockwise closed circular contour centered at the origin and with radius r. The value of r can be chosen as any value for which $X(z)$ converges—i.e., any value such that the circular contour of integration $|z| = r$ is in the ROC. Equation (10.41) is the formal expression for the inverse z-transform and is the discrete-time counterpart of eq. (9.56) for the inverse Laplace transform. As with eq. (9.56), formal evaluation of the inverse transform equation (10.41) requires the use of contour integration in the complex plane. There are, however, a number of alternative procedures for obtaining a sequence from its z-transform. As with Laplace transforms, one particularly useful procedure for rational z-transforms consists of expanding the algebraic expression into a partial-fraction expansion and recognizing the sequences associated with the individual terms. In the following examples, we illustrate the procedure.

Example 10.9

Consider the z-transform

$$X(z) = \frac{3 - \frac{5}{6}z^{-1}}{(1 - \frac{1}{4}z^{-1})(1 - \frac{1}{3}z^{-1})}, \quad |z| > \tfrac{1}{3}. \tag{10.42}$$

There are two poles, one at $z = 1/3$ and one at $z = 1/4$, and the ROC lies outside the outermost pole. That is, the ROC consists of all points with magnitude greater than that of the pole with the larger magnitude, namely the pole at $z = 1/3$. From Property 4 in Section 10.2, we then know that the inverse transform is a right-sided sequence. As described in the appendix, $X(z)$ can be expanded by the method of partial fractions. For this example, the partial-fraction expansion, expressed in polynomials in z^{-1}, is

$$X(z) = \frac{1}{1-\frac{1}{4}z^{-1}} + \frac{2}{1-\frac{1}{3}z^{-1}}. \tag{10.43}$$

Thus, $x[n]$ is the sum of two terms, one with z-transform $1/[1-(1/4)z^{-1}]$ and the other with z-transform $2/[1-(1/3)z^{-1}]$. In order to determine the inverse z-transform of each of these individual terms, we must specify the ROC associated with each. Since the ROC for $X(z)$ is outside the outermost pole, the ROC for each individual term in eq. (10.43) must also be outside the pole associated with that term. That is, the ROC for each term consists of all points with magnitude greater than the magnitude of the corresponding pole. Thus,

$$x[n] = x_1[n] + x_2[n], \tag{10.44}$$

where

$$x_1[n] \overset{z}{\longleftrightarrow} \frac{1}{1-\frac{1}{4}z^{-1}}, \quad |z| > \tfrac{1}{4}, \tag{10.45}$$

$$x_2[n] \overset{z}{\longleftrightarrow} \frac{2}{1-\frac{1}{3}z^{-1}}, \quad |z| > \tfrac{1}{3}. \tag{10.46}$$

From Example 10.1, we can identify by inspection that

$$x_1[n] = \left(\frac{1}{4}\right)^n u[n] \tag{10.47}$$

and

$$x_2[n] = 2\left(\frac{1}{3}\right)^n u[n], \tag{10.48}$$

and thus,

$$x[n] = \left(\frac{1}{4}\right)^n u[n] + 2\left(\frac{1}{3}\right)^n u[n]. \tag{10.49}$$

Example 10.10

Now let us consider the same algebraic expression for $X(z)$ as in eq. (10.42), but with the ROC for $X(z)$ as $1/4 < |z| < 1/3$. Equation (10.43) is still a valid partial-fraction expansion of the algebraic expression for $X(z)$, but the ROC associated with the individual terms will change. In particular, since the ROC for $X(z)$ is outside the pole at $z = 1/4$, the ROC corresponding to this term in eq. (10.43) is also outside the pole and consists of all points with magnitude greater than 1/4, as it did in the previous example. However, since in this example the ROC for $X(z)$ is inside the pole at $z = 1/3$, that is, since the points in the ROC all have magnitude less than 1/3, the ROC corresponding to this term must also lie inside this pole. Thus, the z-transform pairs for the individual components in eq. (10.44) are

$$x_1[n] \overset{z}{\longleftrightarrow} \frac{1}{1-\frac{1}{4}z^{-1}}, \quad |z| > \tfrac{1}{4}, \tag{10.50}$$

and

$$x_2[n] \overset{z}{\longleftrightarrow} \frac{2}{1-\frac{1}{3}z^{-1}}, \quad |z| < \tfrac{1}{3}. \tag{10.51}$$

The signal $x_1[n]$ remains as in eq. (10.47), while from Example 10.2, we can identify

$$x_2[n] = -2\left(\frac{1}{3}\right)^n u[-n-1], \tag{10.52}$$

so that

$$x[n] = \left(\frac{1}{4}\right)^n u[n] - 2\left(\frac{1}{3}\right)^n u[-n-1]. \tag{10.53}$$

Example 10.11

Finally, consider $X(z)$ as in eq. (10.42), but now with the ROC $|z| < 1/4$. In this case the ROC is inside both poles, i.e., the points in the ROC all have magnitude smaller than either of the poles at $z = 1/3$ or $z = 1/4$. Consequently the ROC for each term in the partial-fraction expansion in eq. (10.43) must also lie inside the corresponding pole. As a result, the z-transform pair for $x_1[n]$ is given by

$$x_1[n] \overset{z}{\longleftrightarrow} \frac{1}{1-\frac{1}{4}z^{-1}}, \quad |z| < \frac{1}{4}, \tag{10.54}$$

while the z-transform pair for $x_2[n]$ is given by eq. (10.51). Applying the result of Example 10.2 to eq. (10.54), we find that

$$x_1[n] = -\left(\frac{1}{4}\right)^n u[-n-1],$$

so that

$$x[n] = -\left(\frac{1}{4}\right)^n u[-n-1] - 2\left(\frac{1}{3}\right)^n u[-n-1].$$

The foregoing examples illustrate the basic procedure of using partial-fraction expansions to determine inverse z-transforms. As with the corresponding method for the Laplace transform, the procedure relies on expressing the z-transform as a linear combination of simpler terms. The inverse transform of each term can then be obtained by inspection. In particular, suppose that the partial-fraction expansion of $X(z)$ is of the form

$$X(z) = \sum_{i=1}^{m} \frac{A_i}{1-a_i z^{-1}}, \tag{10.55}$$

so that the inverse transform of $X(z)$ equals the sum of the inverse transforms of the individual terms in the equation. If the ROC of $X(z)$ is outside the pole at $z = a_i$, the inverse transform of the corresponding term in eq. (10.55) is $A_i a_i^n u[n]$. On the other hand, if the ROC of $X(z)$ is inside the pole at $z = a_i$, the inverse transform of this term is $-A_i a_i^n u[-n-1]$. In general, the partial-fraction expansion of a rational transform may include terms in addition to the first-order terms in eq. (10.55). In Section 10.6, we list a number of other z-transform pairs that can be used in conjunction with the z-transform properties to be developed in Section 10.5 to extend the inverse transform method outlined in the preceding example to arbitrary rational z-transforms.

Another very useful procedure for determining the inverse z-transform relies on a power-series expansion of $X(z)$. This procedure is motivated by the observation that the definition of the z-transform given in eq. (10.3) can be interpreted as a power series involving both positive and negative powers of z. The coefficients in this power series are, in fact, the sequence values $x[n]$. To illustrate how a power-series expansion can be used to obtain the inverse z-transform, let us consider three examples.

Example 10.12

Consider the z-transform

$$X(z) = 4z^2 + 2 + 3z^{-1},\ 0 < |z| < \infty. \tag{10.56}$$

From the power-series definition of the z-transform in eq. (10.3), we can determine the inverse transform of $X(z)$ by inspection:

$$x[n] = \begin{cases} 4, & n = -2 \\ 2, & n = 0 \\ 3, & n = 1 \\ 0, & \text{otherwise} \end{cases}.$$

That is,

$$x[n] = 4\delta[n+2] + 2\delta[n] + 3\delta[n-1]. \tag{10.57}$$

Comparing eqs. (10.56) and (10.57), we see that different powers of z serve as placeholders for sequence values at different points in time; i.e., if we simply use the transform pair

$$\delta[n+n_0] \overset{z}{\longleftrightarrow} z^{n_0},$$

we can immediately pass from eq. (10.56) to (10.57) and vice versa.

Example 10.13

Consider

$$X(z) = \frac{1}{1 - az^{-1}}, \quad |z| > |a|.$$

This expression can be expanded in a power series by long division:

$$\begin{array}{r l}
 & 1 + az^{-1} + a^2z^{-2} + \cdots \\ \cline{2-2}
1 - az^{-1} & \big)\, 1 \\
 & 1 - az^{-1} \\ \cline{2-2}
 & \quad az^{-1} \\
 & \quad az^{-1} - a^2z^{-2} \\ \cline{2-2}
 & \qquad\quad a^2z^{-2}
\end{array}$$

or

$$\frac{1}{1-az^{-1}} = 1 + az^{-1} + a^2z^{-2} + \ldots. \tag{10.58}$$

The series expansion of eq. (10.58) converges, since $|z| > |a|$, or equivalently, $|az^{-1}| < 1$. Comparing this equation with the definition of the z-transform in equation (10.3), we see, by matching terms in powers of z, that $x[n] = 0, n < 0$; $x[0] = 1$; $x[1] = a$; $x[2] = a^2$; and in general, $x[n] = a^n u[n]$, which is consistent with Example 10.1.

If, instead, the ROC of $X(z)$ is specified as $|z| < |a|$ or, equivalently, $|az^{-1}| > 1$, then the power-series expansion for $1/(1 - az^{-1})$ in eq. (10.58) does not converge. However, we can obtain a convergent power series by long division again:

$$\begin{array}{r l}
 & -a^{-1}z - a^{-2}z^2 - \cdots \\
\cline{2-2}
-az^{-1} + 1\,\big) & \quad 1 \\
 & \quad 1 - a^{-1}z \\
\cline{2-2}
 & \qquad a^{-1}z
\end{array},$$

or

$$\frac{1}{1-az^{-1}} = -a^{-1}z - a^{-2}z^2 - \ldots. \tag{10.59}$$

In this case, then, $x[n] = 0, n \geq 0$; and $x[-1] = -a^{-1}$, $x[-2] = -a^{-2}, \ldots$; that is, $x[n] = -a^n u[-n-1]$. This is consistent with Example 10.2.

The power-series expansion method for obtaining the inverse z-transform is particularly useful for nonrational z-transforms, which we illustrate with the following example:

Example 10.14

Consider the z-transform

$$X(z) = \log(1 + az^{-1}), \quad |z| > |a|. \tag{10.60}$$

With $|z| > |a|$, or, equivalently, $|az^{-1}| < 1$, eq. (10.60) can be expanded in a power series using the Taylor's series expansion

$$\log(1 + v) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}v^n}{n}, \quad |v| < 1. \tag{10.61}$$

Applying this to eq. (10.60), we have

$$X(z) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}a^n z^{-n}}{n}, \tag{10.62}$$

from which we can identify

$$x[n] = \begin{cases} (-1)^{n+1}\dfrac{a^n}{n}, & n \geq 1 \\ 0, & n \leq 0 \end{cases}, \tag{10.63}$$

or equivalently,

$$x[n] = \frac{-(-a)^n}{n} u[n-1].$$

In Problem 10.63 we consider a related example with region of convergence $|z| < |a|$.

10.4 GEOMETRIC EVALUATION OF THE FOURIER TRANSFORM FROM THE POLE-ZERO PLOT

In Section 10.1 we noted that the z-transform reduces to the Fourier transform for $|z| = 1$ (i.e., for the contour in the z-plane corresponding to the unit circle), provided that the ROC of the z-transform includes the unit circle, so that the Fourier transform converges. In a similar manner, we saw in Chapter 9 that, for continuous-time signals, the Laplace transform reduces to the Fourier transform on the $j\omega$-axis in the s-plane. In Section 9.4, we also discussed the geometric evaluation of the continuous-time Fourier transform from the pole-zero plot. In the discrete-time case, the Fourier transform can again be evaluated geometrically by considering the pole and zero vectors in the z-plane. However, since in this case the rational function is to be evaluated on the contour $|z| = 1$, we consider the vectors from the poles and zeros to the unit circle rather than to the imaginary axis. To illustrate the procedure, let us consider first-order and second-order systems, as discussed in Section 6.6.

10.4.1 First-Order Systems

The impulse response of a first-order causal discrete-time system is of the general form

$$h[n] = a^n u[n], \tag{10.64}$$

and from Example 10.1, its z-transform is

$$H(z) = \frac{1}{1 - az^{-1}} = \frac{z}{z-a}, \quad |z| > |a|. \tag{10.65}$$

For $|a| < 1$, the ROC includes the unit circle, and consequently, the Fourier transform of $h[n]$ converges and is equal to $H(z)$ for $z = e^{j\omega}$. Thus, the frequency response for the first-order system is

$$H(e^{j\omega}) = \frac{1}{1 - ae^{-j\omega}}. \tag{10.66}$$

Figure 10.13(a) depicts the pole-zero plot for $H(z)$ in eq. (10.65), including the vectors from the pole (at $z = a$) and zero (at $z = 0$) to the unit circle. With this plot, the geometric evaluation of $H(z)$ can be carried out using the same procedure as described in Section 9.4. In particular, if we wish to evaluate the frequency response in eq. (10.65), we perform the evaluation for values of z of the form $z = e^{j\omega}$. The magnitude of the frequency response at frequency ω is the ratio of the length of the vector $\mathbf{v}_1$ to the length of the vector $\mathbf{v}_2$ shown in Figure 10.13(a). The phase of the frequency response is the angle of $\mathbf{v}_1$ with respect to the real axis minus the angle of $\mathbf{v}_2$. Furthermore, the vector $\mathbf{v}_1$ from

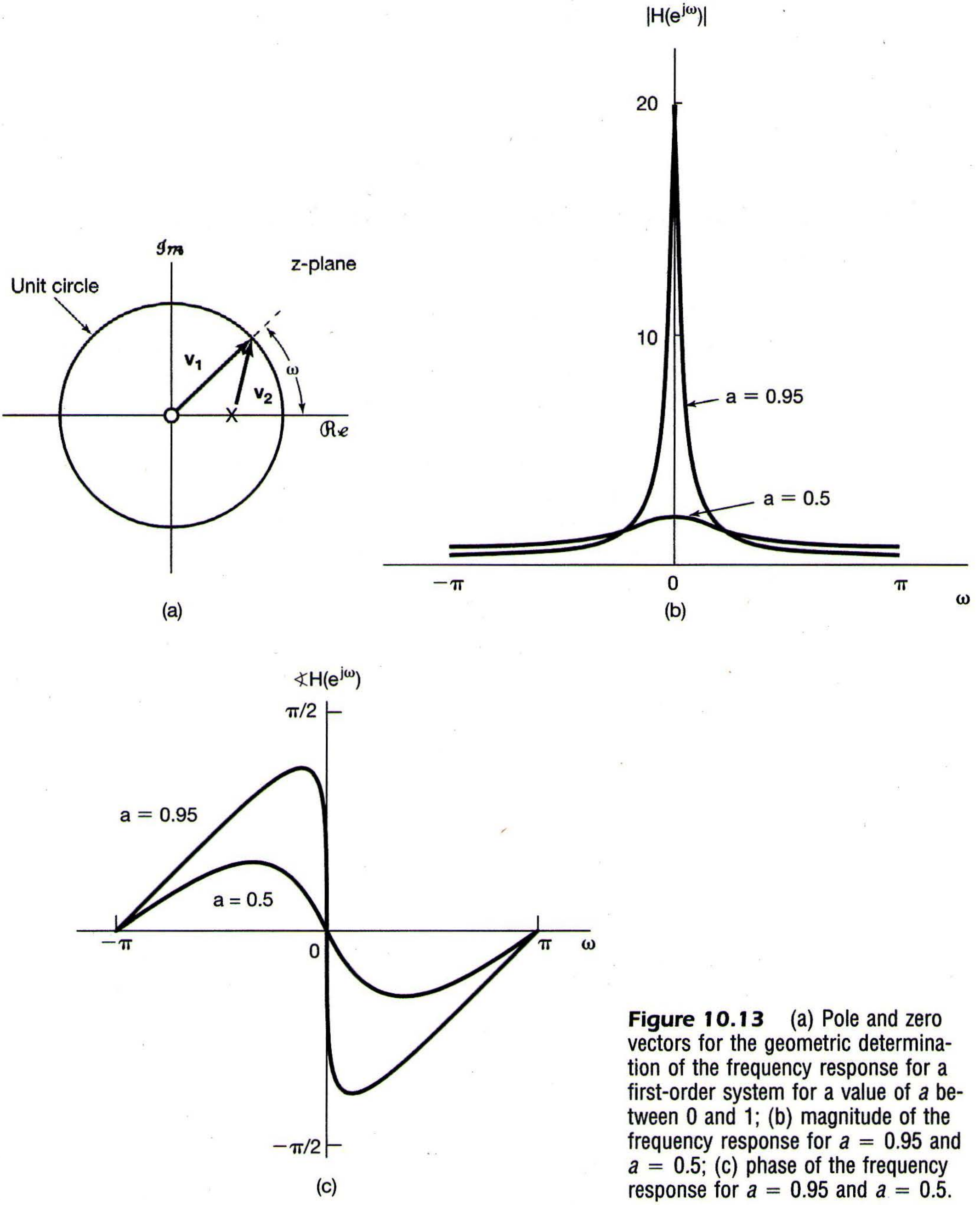


Figure 10.13 (a) Pole and zero vectors for the geometric determination of the frequency response for a first-order system for a value of a between 0 and 1; (b) magnitude of the frequency response for $a = 0.95$ and $a = 0.5$; (c) phase of the frequency response for $a = 0.95$ and $a = 0.5$.

the zero at the origin to the unit circle has a constant length of unity and thus has no effect on the magnitude of $H(e^{j\omega})$. The phase contributed to $H(e^{j\omega})$ by the zero is the angle of the zero vector with respect to the real axis, which we see is equal to ω. For $0 < a < 1$, the pole vector has minimum length at $\omega = 0$ and monotonically increases in length as ω increases from zero to π. Thus, the magnitude of the frequency response will be maximum at $\omega = 0$ and will decrease monotonically as ω increases from 0 to π. The angle of the pole vector begins at zero and increases monotonically as ω increases from zero to π. The resulting magnitude and phase of $H(e^{j\omega})$ are shown in Figures 10.13(b) and (c), respectively, for two values of a.

The magnitude of the parameter a in the discrete-time first-order system plays a role similar to that of the time constant τ for the continuous-time first-order system of Section 9.4.1. Note first that, as illustrated in Figure 10.13, the magnitude of the peak of $H(e^{j\omega})$ at $\omega = 0$ decreases as $|a|$ decreases toward 0. Also, as was discussed in Section 6.6.1 and illustrated in Figures 6.26 and 6.27, as $|a|$ decreases, the impulse response decays more sharply and the step response settles more quickly. With multiple poles, the speed of response associated with each pole is related to its distance from the origin, with those closest to the origin contributing the most rapidly decaying terms in the impulse response. This is further illustrated in the case of second-order systems, which we consider next.

10.4.2 Second-Order Systems

Next, let us consider the class of second-order systems as discussed in Section 6.6.2, with impulse response and frequency response given in eqs. (6.64) and (6.60), which we respectively repeat here as

$$h[n] = r^n \frac{\sin(n+1)\theta}{\sin\theta} u[n] \tag{10.67}$$

and

$$H(e^{j\omega}) = \frac{1}{1 - 2r\cos\theta e^{-j\omega} + r^2 e^{-j2\omega}}, \tag{10.68}$$

where $0 < r < 1$ and $0 \le \theta \le \pi$. Since $H(e^{j\omega}) = H(z)|_{z=e^{j\omega}}$, we can infer from eq. (10.68) that the system function, corresponding to the z-transform of the system impulse response, is

$$H(z) = \frac{1}{1 - (2r\cos\theta)z^{-1} + r^2 z^{-2}}. \tag{10.69}$$

The poles of $H(z)$ are located at

$$z_1 = re^{j\theta}, \quad z_2 = re^{-j\theta}, \tag{10.70}$$

and there is a double zero at $z = 0$. The pole-zero plot and the pole and zero vectors with $0 < \theta < \pi/2$ are illustrated in Figure 10.14(a). In this case, the magnitude of the frequency response equals the square of the magnitude of $\mathbf{v}_1$ (since there is a double zero at the origin) divided by the product of the magnitudes of $\mathbf{v}_2$ and $\mathbf{v}_3$. Because the length of the vector $\mathbf{v}_1$ from the zero at the origin is 1 for all values of ω, the magnitude of the frequency response equals the reciprocal of the product of the lengths of the two pole vectors $\mathbf{v}_2$ and $\mathbf{v}_3$. Also, the phase of the frequency response equals twice the angle of $\mathbf{v}_1$ with respect to the real axis minus the sums of the angles of $\mathbf{v}_2$ and $\mathbf{v}_3$. In Figure 10.14(b) we show the magnitude of the frequency response for $r = 0.95$ and $r = 0.75$, while in Figure 10.14(c) we display the phase of $H(e^{j\omega})$ for the same two values of r. We note in particular that, as we move

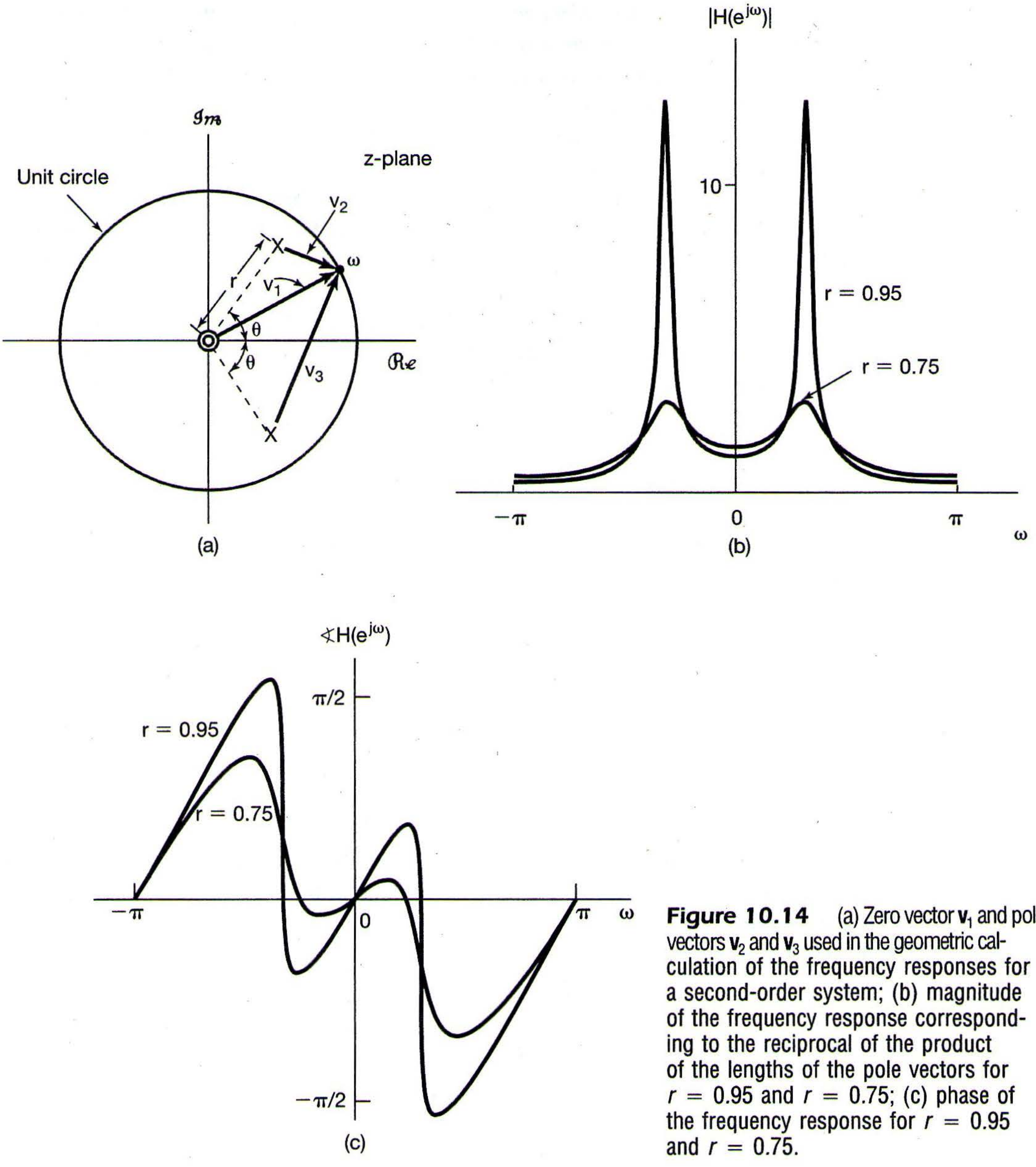


Figure 10.14 (a) Zero vector $\mathbf{v}_1$ and pole vectors $\mathbf{v}_2$ and $\mathbf{v}_3$ used in the geometric calculation of the frequency responses for a second-order system; (b) magnitude of the frequency response corresponding to the reciprocal of the product of the lengths of the pole vectors for $r = 0.95$ and $r = 0.75$; (c) phase of the frequency response for $r = 0.95$ and $r = 0.75$.

along the unit circle from $\omega = 0$ toward $\omega = \pi$, the length of the vector $\mathbf{v}_2$ first decreases and then increases, with a minimum length in the vicinity of the pole location, at $\omega = \theta$. This is consistent with the fact that the magnitude of the frequency response peaks for ω near θ when the length of the vector $\mathbf{v}_2$ is small. Based on the behavior of the pole vectors, it is also evident that as r increases toward unity, the minimum length of the pole vectors will decrease, causing the frequency response to peak more sharply with increasing r. Also, for r near unity, the angle of the vector $\mathbf{v}_2$ changes sharply for ω in the vicinity of θ. Furthermore, from the form of the impulse response [eq. (10.67) and Figure 6.29] or the step response [eq. (6.67) and Figure 6.30], we see, as we did with the first-order system, that as the poles move closer to the origin, corresponding to r decreasing, the impulse response decays more rapidly and the step response settles more quickly.

10.5 PROPERTIES OF THE z-TRANSFORM

As with the other transforms we have developed, the z-transform possesses a number of properties that make it an extremely valuable tool in the study of discrete-time signals and systems. In this section, we summarize many of these properties. Their derivations are analogous to the derivations of properties for the other transforms, and thus, many are left as exercises at the end of the chapter. (See Problems 10.43 and 10.51–10.54.)

10.5.1 Linearity

If

$$x_1[n] \overset{z}{\longleftrightarrow} X_1(z), \quad \text{with ROC} = R_1,$$

and

$$x_2[n] \overset{z}{\longleftrightarrow} X_2(z), \quad \text{with ROC} = R_2,$$

then

$$\boxed{ax_1[n] + bx_2[n] \overset{z}{\longleftrightarrow} aX_1(z) + bX_2(z), \quad \text{with ROC containing } R_1 \cap R_2.} \tag{10.71}$$

As indicated, the ROC of the linear combination is at least the intersection of R_1 and R_2. For sequences with rational z-transforms, if the poles of $aX_1(z) + bX_2(z)$ consist of all of the poles of $X_1(z)$ and $X_2(z)$ (i.e., if there is no pole-zero cancellation), then the region of convergence will be exactly equal to the overlap of the individual regions of convergence. If the linear combination is such that some zeros are introduced that cancel poles, then the region of convergence may be larger. A simple example of this occurs when $x_1[n]$ and $x_2[n]$ are both of infinite duration, but the linear combination is of finite duration. In this case the region of convergence of the linear combination is the entire z-plane, with the possible exception of zero and/or infinity. For example, the sequences $a^n u[n]$ and $a^n u[n-1]$ both have a region of convergence defined by $|z| > |a|$, but the sequence corresponding to the difference $(a^n u[n] - a^n u[n-1]) = \delta[n]$ has a region of convergence that is the entire z-plane.

10.5.2 Time Shifting

If

$$x[n] \overset{z}{\longleftrightarrow} X(z), \quad \text{with ROC} = R,$$

then

$$\boxed{x[n-n_0] \overset{\mathcal{Z}}{\longleftrightarrow} z^{-n_0}X(z), \quad \text{with ROC} = R\text{, except for the possible addition or deletion of the origin or infinity.}} \tag{10.72}$$

Because of the multiplication by z^{-n_0}, for $n_0 > 0$ poles will be introduced at $z = 0$, which may cancel corresponding zeros of $X(z)$ at $z = 0$. Consequently, $z = 0$ may be a pole of $z^{-n_0}X(z)$ while it may not be a pole of $X(z)$. In this case the ROC for $z^{-n_0}X(z)$ equals the ROC of $X(z)$ but with the origin deleted. Similarly, if $n_0 < 0$, zeros will be introduced at $z = 0$, which may cancel corresponding poles of $X(z)$ at $z = 0$. Consequently, $z = 0$ may be a zero of $z^{-n_0}X(z)$ while it may not be a pole of $X(z)$. In this case $z = \infty$ is a pole of $z^{-n_0}X(z)$, and thus the ROC for $z^{-n_0}X(z)$ equals the ROC of $X(z)$ but with the $z = \infty$ deleted.

10.5.3 Scaling in the z-Domain

If

$$x[n] \overset{\mathcal{Z}}{\longleftrightarrow} X(z), \quad \text{with ROC} = R,$$

then

$$\boxed{z_0^n x[n] \overset{\mathcal{Z}}{\longleftrightarrow} X\left(\frac{z}{z_0}\right), \quad \text{with ROC} = |z_0|R,} \tag{10.73}$$

where $|z_0|R$ is the scaled version of R. That is, if z is a point in the ROC of $X(z)$, then the point $|z_0|z$ is in the ROC of $X(z/z_0)$. Also, if $X(z)$ has a pole (or zero) at $z = a$, then $X(z/z_0)$ has a pole (or zero) at $z = z_0 a$.

An important special case of eq. (10.73) is when $z_0 = e^{j\omega_0}$. In this case, $|z_0|R = R$ and

$$e^{j\omega_0 n}x[n] \overset{\mathcal{Z}}{\longleftrightarrow} X(e^{-j\omega_0}z). \tag{10.74}$$

The left-hand side of eq. (10.74) corresponds to multiplication by a complex exponential sequence. The right-hand side can be interpreted as a rotation in the z-plane; that is, all pole-zero locations rotate in the z-plane by an angle of ω_0, as illustrated in Figure 10.15. This can be seen by noting that if $X(z)$ has a factor of the form $1 - az^{-1}$, then $X(e^{-j\omega_0}z)$ will have a factor $1 - ae^{j\omega_0}z^{-1}$, and thus, a pole or zero at $z = a$ in $X(z)$ will become a pole or zero at $z = ae^{j\omega_0}$ in $X(e^{-j\omega_0}z)$. The behavior of the z-transform on the unit circle will then also shift by an angle of ω_0. This is consistent with the frequency-shifting property set forth in Section 5.3.3, where multiplication with a complex exponential in the time domain was shown to correspond to a shift in frequency of the Fourier transform. Also, in the more general case when $z_0 = r_0 e^{j\omega_0}$ in eq. (10.73), the pole and zero locations are rotated by ω_0 and scaled in magnitude by a factor of r_0.

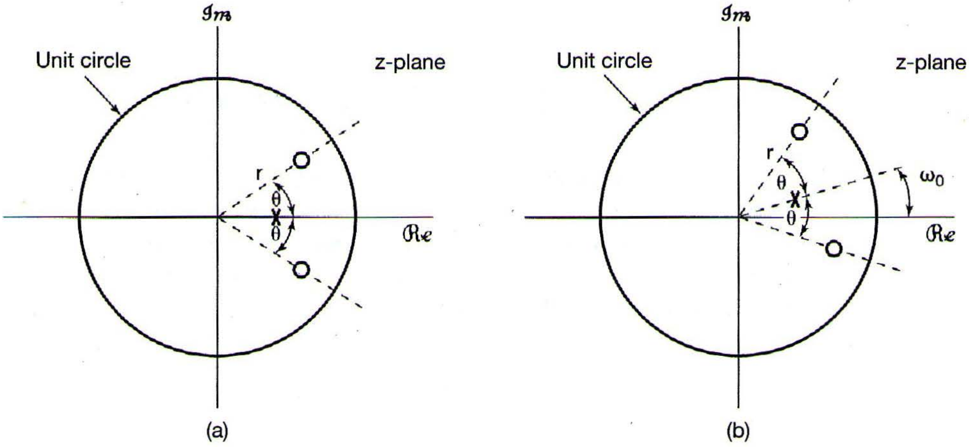


Figure 10.15 Effect on the pole-zero plot of time-domain multiplication by a complex exponential sequence $e^{j\omega_0 n}$: (a) pole-zero pattern for the z-transform for a signal $x[n]$; (b) pole-zero pattern for the z-transform of $x[n]e^{j\omega_0 n}$.

10.5.4 Time Reversal

If

$$x[n] \overset{z}{\longleftrightarrow} X(z), \quad \text{with ROC} = R,$$

then

$$\boxed{x[-n] \overset{z}{\longleftrightarrow} X(\tfrac{1}{z}), \quad \text{with ROC} = \tfrac{1}{R}.} \tag{10.75}$$

That is, if z_0 is in the ROC for $x[n]$, then $1/z_0$ is in the ROC for $x[-n]$.

10.5.5 Time Expansion

As we discussed in Section 5.3.7, the continuous-time concept of time scaling does not directly extend to discrete time, since the discrete-time index is defined only for integer values. However, the discrete-time concept of time expansion—i.e., of inserting a number of zeros between successive values of a discrete-time sequence $x[n]$—can be defined and does play an important role in discrete-time signal and system analysis. Specifically, the sequence $x_{(k)}[n]$, introduced in Section 5.3.7 and defined as

$$x_{(k)}[n] = \begin{cases} x[n/k], & \text{if } n \text{ is a multiple of } k \\ 0, & \text{if } n \text{ is not a multiple of } k \end{cases} \tag{10.76}$$

has $k-1$ zeros inserted between successive values of the original signal. In this case, if

$$x[n] \overset{z}{\longleftrightarrow} X(z), \quad \text{with ROC} = R,$$

then

$$\boxed{x_{(k)}[n] \overset{z}{\longleftrightarrow} X(z^k), \quad \text{with ROC} = R^{1/k}.} \tag{10.77}$$

That is, if z is in the ROC of $X(z)$, then the point $z^{1/k}$ is in the ROC of $X(z^k)$. Also, if $X(z)$ has a pole (or zero) at $z = a$, then $X(z^k)$ has a pole (or zero) at $z = a^{1/k}$.

The interpretation of this result follows from the power-series form of the z-transform, from which we see that the coefficient of the term z^{-n} equals the value of the signal at time n. That is, with

$$X(z) = \sum_{n=-\infty}^{+\infty} x[n]z^{-n},$$

it follows that

$$X(z^k) = \sum_{n=-\infty}^{+\infty} x[n](z^k)^{-n} = \sum_{n=-\infty}^{+\infty} x[n]z^{-kn}. \tag{10.78}$$

Examining the right-hand side of eq. (10.78), we see that the only terms that appear are of the form z^{-kn}. In other words, the coefficient of the term z^{-m} in this power series equals 0 if m is not a multiple of k and equals $x[m/k]$ if m is a multiple of k. Thus, the inverse transform of eq. (10.78) is $x_{(k)}[n]$.

10.5.6 Conjugation

If

$$x[n] \overset{z}{\longleftrightarrow} X(z), \quad \text{with ROC} = R, \tag{10.79}$$

then

$$\boxed{x^*[n] \overset{z}{\longleftrightarrow} X^*(z^*), \quad \text{with ROC} = R.} \tag{10.80}$$

Consequently, if $x[n]$ is real, we can conclude from eq. (10.80) that

$$X(z) = X^*(z^*).$$

Thus, if $X(z)$ has a pole (or zero) at $z = z_0$, it must also have a pole (or zero) at the complex conjugate point $z = z_0^*$. For example, the transform $X(z)$ for the real signal $x[n]$ in Example 10.4 has poles at $z = (1/3)e^{\pm j\pi/4}$.

10.5.7 The Convolution Property

If

$$x_1[n] \overset{z}{\longleftrightarrow} X_1(z), \quad \text{with ROC} = R_1,$$

and

$$x_2[n] \overset{z}{\longleftrightarrow} X_2(z), \quad \text{with ROC} = R_2,$$

then

$$\boxed{x_1[n] * x_2[n] \overset{z}{\longleftrightarrow} X_1(z)X_2(z), \quad \text{with ROC containing } R_1 \cap R_2.} \tag{10.81}$$

Just as with the convolution property for the Laplace transform, the ROC of $X_1(z)X_2(z)$ includes the intersection of R_1 and R_2 and may be larger if pole-zero cancellation occurs in the product. The convolution property for the z-transform can be derived in a variety of different ways. A formal derivation is developed in Problem 10.56. A derivation can also be carried out analogous to that used for the convolution property for the continuous-time Fourier transform in Section 4.4, which relied on the interpretation of the Fourier transform as the change in amplitude of a complex exponential through an LTI system.

For the z-transform, there is another often useful interpretation of the convolution property. From the definition in eq. (10.3), we recognize the z-transform as a series in z^{-1} where the coefficient of z^{-n} is the sequence value $x[n]$. In essence, the convolution property equation (10.81) states that when two polynomials or power series $X_1(z)$ and $X_2(z)$ are multiplied, the coefficients in the polynomial representing the product are the convolution of the coefficients in the polynomials $X_1(z)$ and $X_2(z)$. (See Problem 10.57).

Example 10.15

Consider an LTI system for which

$$y[n] = h[n] * x[n], \tag{10.82}$$

where

$$h[n] = \delta[n] - \delta[n-1].$$

Note that

$$\delta[n] - \delta[n-1] \overset{z}{\longleftrightarrow} 1 - z^{-1}, \tag{10.83}$$

with ROC equal to the entire z-plane except the origin. Also, the z-transform in eq. (10.83) has a zero at $z = 1$. From eq. (10.81), we see that if

$$x[n] \overset{z}{\longleftrightarrow} X(z), \quad \text{with ROC} = R,$$

then

$$y[n] \overset{z}{\longleftrightarrow} (1 - z^{-1})X(z), \tag{10.84}$$

with ROC equal to R, with the possible deletion of $z = 0$ and/or addition of $z = 1$.

Note that for this system

$$y[n] = [\delta[n] - \delta[n-1]] * x[n] = x[n] - x[n-1].$$

That is, $y[n]$ is the first difference of the sequence $x[n]$. Since the first-difference operation is commonly thought of as a discrete-time counterpart to differentiation, eq. (10.83) can be thought of as the z-transform counterpart of the Laplace transform differentiation property presented in Section 9.5.7.

Example 10.16

Suppose we now consider the inverse of first differencing, namely, accumulation or summation. Specifically, let $w[n]$ be the running sum of $x[n]$:

$$w[n] = \sum_{k=-\infty}^{n} x[k] = u[n] * x[n]. \tag{10.85}$$

Then, using eq. (10.81) together with the z-transform of the unit step in Example 10.1, we see that

$$w[n] = \sum_{k=-\infty}^{n} x[k] \overset{z}{\longleftrightarrow} \frac{1}{1-z^{-1}}X(z), \tag{10.86}$$

with ROC including at least the intersection of R with $|z| > 1$. Eq. (10.86) is the discrete-time z-transform counterpart of the integration property in Section 9.5.9.

10.5.8 Differentiation in the z-Domain

If

$$x[n] \overset{z}{\longleftrightarrow} X(z), \quad \text{with ROC} = R,$$

then

$$\boxed{nx[n] \overset{z}{\longleftrightarrow} -z\frac{dX(z)}{dz}, \quad \text{with ROC} = R.} \tag{10.87}$$

This property follows in a straightforward manner by differentiating both sides of the expression for the z-transform given in eq. (10.3). As an example of the use of this property, let us apply it to determining the inverse z-transform considered in Example 10.14.

Example 10.17

If

$$X(z) = \log(1 + az^{-1}), \quad |z| > |a|, \tag{10.88}$$

then

$$nx[n] \overset{z}{\longleftrightarrow} -z\frac{dX(z)}{dz} = \frac{az^{-1}}{1+az^{-1}}, \quad |z| > |a|. \tag{10.89}$$

By differentiating, we have converted the z-transform to a rational expression. The inverse z-transform of the right-hand side of eq. (10.89) can be obtained by using Example 10.1 together with the time-shifting property, eq. (10.72), set forth in Section 10.5.2.

Specifically, from Example 10.1 and the linearity property,

$$a(-a)^n u[n] \overset{z}{\longleftrightarrow} \frac{a}{1+az^{-1}}, \quad |z| > |a|. \tag{10.90}$$

Combining this with the time-shifting property yields

$$a(-a)^{n-1} u[n-1] \overset{z}{\longleftrightarrow} \frac{az^{-1}}{1+az^{-1}}, \quad |z| > |a|.$$

Consequently,

$$x[n] = \frac{-(-a)^n}{n} u[n-1]. \tag{10.91}$$

Example 10.18

As another example of the use of the differentiation property, consider determining the inverse z-transform for

$$X(z) = \frac{az^{-1}}{(1-az^{-1})^2}, \quad |z| > |a|. \tag{10.92}$$

From Example 10.1,

$$a^n u[n] \overset{z}{\longleftrightarrow} \frac{1}{1-az^{-1}}, \quad |z| > |a|, \tag{10.93}$$

and hence,

$$na^n u[n] \overset{z}{\longleftrightarrow} -z\frac{d}{dz}\left(\frac{1}{1-az^{-1}}\right) = \frac{az^{-1}}{(1-az^{-1})^2}, \quad |z| > |a|. \tag{10.94}$$

10.5.9 The Initial-Value Theorem

If $x[n] = 0, n < 0$, then

$$x[0] = \lim_{z\to\infty} X(z). \tag{10.95}$$

This property follows by considering the limit of each term individually in the expression for the z-transform, with $x[n]$ zero for $n < 0$. With this constraint,

$$X(z) = \sum_{n=0}^{\infty} x[n]z^{-n}.$$

As $z \to \infty$, $z^{-n} \to 0$ for $n > 0$, whereas for $n = 0$, $z^{-n} = 1$. Thus, eq. (10.95) follows.

As one consequence of the initial-value theorem, for a causal sequence, if $x[0]$ is finite, then $\lim_{z\to\infty} X(z)$ is finite. Consequently, with $X(z)$ expressed as a ratio of polynomials in z, the order of the numerator polynomial cannot be greater than the order of the denominator polynomial; or, equivalently, the number of finite zeros of $X(z)$ cannot be greater than the number of finite poles.

Example 10.19

The initial-value theorem can also be useful in checking the correctness of the z-transform calculation for a signal. For example, consider the signal $x[n]$ in Example 10.3. From eq. (10.12), we see that $x[0] = 1$. Also, from eq. (10.14),

$$\lim_{z\to\infty} X(z) = \lim_{z\to\infty} \frac{1-\frac{3}{2}z^{-1}}{(1-\frac{1}{3}z^{-1})(1-\frac{1}{2}z^{-1})} = 1,$$

which is consistent with the initial-value theorem.

10.5.10 Summary of Properties

In Table 10.1, we summarize the properties of the z-transform.

10.6 SOME COMMON z-TRANSFORM PAIRS

As with the inverse Laplace transform, the inverse z-transform can often be easily evaluated by expressing $X(z)$ as a linear combination of simpler terms, the inverse transforms of which are recognizable. In Table 10.2, we have listed a number of useful z-transform pairs. Each of these can be developed from previous examples in combination with the properties of the z-transform listed in Table 10.1. For example, transform pairs 2 and 5 follow directly from Example 10.1, and transform pair 7 is developed in Example 10.18. These, together with the time-reversal and time-shifting properties set forth in Sections 10.5.4 and 10.5.2, respectively, then lead to transform pairs 3, 6, and 8. Transform pairs 9 and 10 can be developed using transform pair 2 together with the linearity and scaling properties developed in Sections 10.5.1 and 10.5.3, respectively.

10.7 ANALYSIS AND CHARACTERIZATION OF LTI SYSTEMS USING z-TRANSFORMS

The z-transform plays a particularly important role in the analysis and representation of discrete-time LTI systems. From the convolution property presented in Section 10.5.7,

$$Y(z) = H(z)X(z), \tag{10.96}$$

where $X(z)$, $Y(z)$, and $H(z)$ are the z-transforms of the system input, output, and impulse response, respectively. $H(z)$ is referred to as the *system function* or *transfer function* of the system. For z evaluated on the unit circle (i.e., for $z = e^{j\omega}$), $H(z)$ reduces to the frequency response of the system, provided that the unit circle is in the ROC for $H(z)$. Also, from our discussion in Section 3.2, we know that if the input to an LTI system is the complex exponential signal $x[n] = z^n$, then the output will be $H(z)z^n$. That is, z^n is an eigenfunction of the system with eigenvalue given by $H(z)$, the z-transform of the impulse response.

Many properties of a system can be tied directly to characteristics of the poles, zeros, and region of convergence of the system function, and in this section we illustrate some of these relationships by examining several important system properties and an important class of systems.

TABLE 10.1 PROPERTIES OF THE z-TRANSFORM

Section	Property	Signal	z-Transform	ROC
		$x[n]$ $x_1[n]$ $x_2[n]$	$X(z)$ $X_1(z)$ $X_2(z)$	R R_1 R_2
10.5.1	Linearity	$ax_1[n] + bx_2[n]$	$aX_1(z) + bX_2(z)$	At least the intersection of R_1 and R_2
10.5.2	Time shifting	$x[n - n_0]$	$z^{-n_0}X(z)$	R, except for the possible addition or deletion of the origin
10.5.3	Scaling in the z-domain	$e^{j\omega_0 n}x[n]$	$X(e^{-j\omega_0}z)$	R
		$z_0^n x[n]$	$X\left(\frac{z}{z_0}\right)$	$z_0 R$
		$a^n x[n]$	$X(a^{-1}z)$	Scaled version of R (i.e., $\lvert a\rvert R$ = the set of points $\{\lvert a\rvert z\}$ for z in R)
10.5.4	Time reversal	$x[-n]$	$X(z^{-1})$	Inverted R (i.e., R^{-1} = the set of points z^{-1}, where z is in R)
10.5.5	Time expansion	$x_{(k)}[n] = \begin{cases} x[r], & n = rk \\ 0, & n \neq rk \end{cases}$ for some integer r	$X(z^k)$	$R^{1/k}$ (i.e., the set of points $z^{1/k}$, where z is in R)
10.5.6	Conjugation	$x^*[n]$	$X^*(z^*)$	R
10.5.7	Convolution	$x_1[n] * x_2[n]$	$X_1(z)X_2(z)$	At least the intersection of R_1 and R_2
10.5.7	First difference	$x[n] - x[n-1]$	$(1 - z^{-1})X(z)$	At least the intersection of R and $\lvert z\rvert > 0$
10.5.7	Accumulation	$\sum_{k=-\infty}^{n} x[k]$	$\frac{1}{1 - z^{-1}}X(z)$	At least the intersection of R and $\lvert z\rvert > 1$
10.5.8	Differentiation in the z-domain	$nx[n]$	$-z\frac{dX(z)}{dz}$	R
10.5.9		Initial Value Theorem If $x[n] = 0$ for $n < 0$, then $x[0] = \lim_{z\to\infty} X(z)$		

TABLE 10.2 SOME COMMON z-TRANSFORM PAIRS

Signal	Transform	ROC
1. $\delta[n]$	1	All z
2. $u[n]$	$\dfrac{1}{1-z^{-1}}$	$\lvert z\rvert > 1$
3. $-u[-n-1]$	$\dfrac{1}{1-z^{-1}}$	$\lvert z\rvert < 1$
4. $\delta[n-m]$	z^{-m}	All z, except 0 (if $m > 0$) or ∞ (if $m < 0$)
5. $\alpha^n u[n]$	$\dfrac{1}{1-\alpha z^{-1}}$	$\lvert z\rvert > \lvert\alpha\rvert$
6. $-\alpha^n u[-n-1]$	$\dfrac{1}{1-\alpha z^{-1}}$	$\lvert z\rvert < \lvert\alpha\rvert$
7. $n\alpha^n u[n]$	$\dfrac{\alpha z^{-1}}{(1-\alpha z^{-1})^2}$	$\lvert z\rvert > \lvert\alpha\rvert$
8. $-n\alpha^n u[-n-1]$	$\dfrac{\alpha z^{-1}}{(1-\alpha z^{-1})^2}$	$\lvert z\rvert < \lvert\alpha\rvert$
9. $[\cos\omega_0 n]u[n]$	$\dfrac{1-[\cos\omega_0]z^{-1}}{1-[2\cos\omega_0]z^{-1}+z^{-2}}$	$\lvert z\rvert > 1$
10. $[\sin\omega_0 n]u[n]$	$\dfrac{[\sin\omega_0]z^{-1}}{1-[2\cos\omega_0]z^{-1}+z^{-2}}$	$\lvert z\rvert > 1$
11. $[r^n\cos\omega_0 n]u[n]$	$\dfrac{1-[r\cos\omega_0]z^{-1}}{1-[2r\cos\omega_0]z^{-1}+r^2z^{-2}}$	$\lvert z\rvert > r$
12. $[r^n\sin\omega_0 n]u[n]$	$\dfrac{[r\sin\omega_0]z^{-1}}{1-[2r\cos\omega_0]z^{-1}+r^2z^{-2}}$	$\lvert z\rvert > r$

10.7.1 Causality

A causal LTI system has an impulse response $h[n]$ that is zero for $n < 0$, and therefore is right-sided. From Property 4 in Section 10.2 we then know that the ROC of $H(z)$ is the exterior of a circle in the z-plane. For some systems, e.g., if $h[n] = \delta[n]$, so that $H(z) = 1$, the ROC can extend all the way in to and possibly include the origin. Also, in general, for a right-sided impulse response, the ROC may or may not include infinity. For example, if $h[n] = \delta[n+1]$, then $H(z) = z$, which has a pole at infinity. However, as we saw in Property 8 in Section 10.2, for a causal system the power series

$$H(z) = \sum_{n=0}^{\infty} h[n]z^{-n}$$

does not include any positive powers of z. Consequently, the ROC includes infinity. Summarizing, we have the follow principle:

A discrete-time LTI system is causal if and only if the ROC of its system function is the exterior of a circle, including infinity.

If $H(z)$ is rational, then, from Property 8 in Section 10.2, for the system to be causal, the ROC must be outside the outermost pole and infinity must be in the ROC. Equivalently, the limit of $H(z)$ as $z \longrightarrow \infty$ must be finite. As we discussed in Section 10.5.9, this is equivalent to the numerator of $H(z)$ having degree no larger than the denominator when both are expressed as polynomials in z. That is:

A discrete-time LTI system with rational system function $H(z)$ is causal if and only if: (a) the ROC is the exterior of a circle outside the outermost pole; and (b) with $H(z)$ expressed as a ratio of polynomials in z, the order of the numerator cannot be greater than the order of the denominator.

Example 10.20

Consider a system with system function whose algebraic expression is

$$H(z) = \frac{z^3 - 2z^2 + z}{z^2 + \frac{1}{4}z + \frac{1}{8}}.$$

Without even knowing the ROC for this system, we can conclude that the system is not causal, because the numerator of $H(z)$ is of higher order than the denominator.

Example 10.21

Consider a system with system function

$$H(z) = \frac{1}{1 - \frac{1}{2}z^{-1}} + \frac{1}{1 - 2z^{-1}}, \quad |z| > 2 \tag{10.97}$$

Since the ROC for this system function is the exterior of a circle outside the outermost pole, we know that the impulse response is right-sided. To determine if the system is causal, we then need only check the other condition required for causality, namely that $H(z)$, when expressed as a ratio of polynomials in z, has numerator degree no larger than the denominator. For this example,

$$H(z) = \frac{2 - \frac{5}{2}z^{-1}}{(1 - \frac{1}{2}z^{-1})(1 - 2z^{-1})} = \frac{2z^2 - \frac{5}{2}z}{z^2 - \frac{5}{2}z + 1}, \tag{10.98}$$

so that the numerator and denominator of $H(z)$ are both of degree two, and consequently we can conclude that the system is causal. This can also be verified by calculating the inverse transform of $H(z)$. In particular, using transform pair 5 in Table 10.2, we find that the impulse response of this system is

$$h[n] = \left[\left(\frac{1}{2}\right)^n + 2^n\right]u[n]. \tag{10.99}$$

Since $h[n] = 0$ for $n < 0$, we can confirm that the system is causal.

10.7.2 Stability

As we discussed in Section 2.3.7, the stability of a discrete-time LTI system is equivalent to its impulse response being absolutely summable. In this case the Fourier transform of $h[n]$

converges, and consequently, the ROC of $H(z)$ must include the unit circle. Summarizing, we obtain the following result:

> An LTI system is stable if and only if the ROC of its system function $H(z)$ includes the unit circle, $|z| = 1$.

Example 10.22

Consider again the system function in eq. (10.97). Since the associated ROC is the region $|z| > 2$, which does not include the unit circle, the system is not stable. This can also be seen by noting that the impulse response in eq. (10.99) is not absolutely summable. If, however, we consider a system whose system function has the same algebraic expression as in eq. (10.97) but whose ROC is the region $1/2 < |z| < 2$, then the ROC does contain the unit circle, so that the corresponding system is noncausal but stable. In this case, using transform pairs 5 and 6 from Table 10.2, we find that the corresponding impulse response is

$$h[n] = \left(\frac{1}{2}\right)^n u[n] - 2^n u[-n-1], \tag{10.100}$$

which is absolutely summable.

Also, for the third possible choice of ROC associated with the algebraic expression for $H(z)$ in eq. (10.97), namely, $|z| < 1/2$, the corresponding system is neither causal (since the ROC is not outside the outermost pole) nor stable (since the ROC does not include the unit circle). This can also be seen from the impulse response, which (using transform pair 6 in Table 10.2) is

$$h[n] = -\left[\left(\frac{1}{2}\right)^n + 2^n\right] u[-n-1].$$

As Example 10.22 illustrates, it is perfectly possible for a system to be stable but not causal. However, if we focus on causal systems, stability can easily be checked by examining the locations of the poles. Specifically, for a causal system with rational system function, the ROC is outside the outermost pole. For this ROC to include the unit circle, $|z| = 1$, all of the poles of the system must be inside the unit circle. That is:

> A causal LTI system with rational system function $H(z)$ is stable if and only if all of the poles of $H(z)$ lie inside the unit circle—i.e., they must all have magnitude smaller than 1.

Example 10.23

Consider a causal system with system function

$$H(z) = \frac{1}{1 - az^{-1}},$$

which has a pole at $z = a$. For this system to be stable, its pole must be inside the unit circle, i.e., we must have $|a| < 1$. This is consistent with the condition for the absolute summability of the corresponding impulse response $h[n] = a^n u[n]$.

Example 10.24

The system function for a second-order system with complex poles was given in eq. (10.69), specifically,

$$H(z) = \frac{1}{1 - (2r\cos\theta)z^{-1} + r^2z^{-2}}, \tag{10.101}$$

with poles located at $z_1 = re^{j\theta}$ and $z_2 = re^{-j\theta}$. Assuming causality, we see that the ROC is outside the outermost pole (i.e., $|z| > |r|$). The pole-zero plot and ROC for this system are shown in Figure 10.16 for $r < 1$ and $r > 1$. For $r < 1$, the poles are inside the unit circle, the ROC includes the unit circle, and therefore, the system is stable. For $r > 1$, the poles are outside the unit circle, the ROC does not include the unit circle, and the system is unstable.

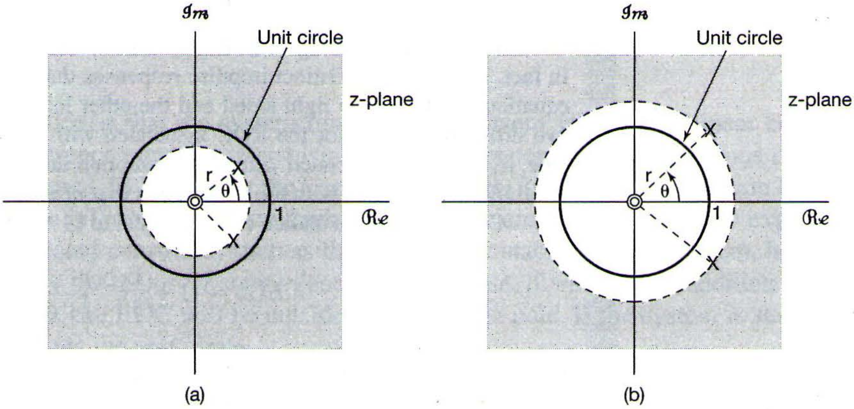


Figure 10.16 Pole-zero plot for a second-order system with complex poles: (a) $r < 1$; (b) $r > 1$.

10.7.3 LTI Systems Characterized by Linear Constant-Coefficient Difference Equations

For systems characterized by linear constant-coefficient difference equations, the properties of the z-transform provide a particularly convenient procedure for obtaining the system function, frequency response, or time-domain response of the system. Let us illustrate this with an example.

Example 10.25

Consider an LTI system for which the input $x[n]$ and output $y[n]$ satisfy the linear constant-coefficient difference equation

$$y[n] - \frac{1}{2}y[n-1] = x[n] + \frac{1}{3}x[n-1]. \tag{10.102}$$

Applying the z-transform to both sides of eq. (10.102), and using the linearity property set forth in Section 10.5.1 and the time-shifting property presented in Section 10.5.2, we obtain

$$Y(z) - \frac{1}{2}z^{-1}Y(z) = X(z) + \frac{1}{3}z^{-1}X(z),$$

or

$$Y(z) = X(z)\left[\frac{1 + \frac{1}{3}z^{-1}}{1 - \frac{1}{2}z^{-1}}\right]. \tag{10.103}$$

From eq. (10.96), then,

$$H(z) = \frac{Y(z)}{X(z)} = \frac{1 + \frac{1}{3}z^{-1}}{1 - \frac{1}{2}z^{-1}}. \tag{10.104}$$

This provides the algebraic expression for $H(z)$, but not the region of convergence. In fact, there are two distinct impulse responses that are consistent with the difference equation (10.102), one right sided and the other left sided. Correspondingly, there are two different choices for the ROC associated with the algebraic expression (10.104). One, $|z| > 1/2$, is associated with the assumption that $h[n]$ is right sided, and the other, $|z| < 1/2$, is associated with the assumption that $h[n]$ is left sided.

Consider first the choice of ROC equal to $|z| > 1$. Writing

$$H(z) = \left(1 + \frac{1}{3}z^{-1}\right)\frac{1}{1 - \frac{1}{2}z^{-1}},$$

we can use transform pair 5 in Table 10.2, together with the linearity and time-shifting properties, to find the corresponding impulse response

$$h[n] = \left(\frac{1}{2}\right)^{n} u[n] + \frac{1}{3}\left(\frac{1}{2}\right)^{n-1} u[n-1].$$

For the other choice of ROC, namely, $|z| < 1$, we can use transform pair 6 in Table 10.2 and the linearity and time-shifting properties, yielding

$$h[n] = -\left(\frac{1}{2}\right)^{n} u[-n-1] - \frac{1}{3}\left(\frac{1}{2}\right)^{n-1} u[-n].$$

In this case, the system is anticausal ($h[n] = 0$ for $n > 0$) and unstable.

For the more general case of an Nth-order difference equation, we proceed in a manner similar to that in Example 10.25, applying the z-transform to both sides of the equation and using the linearity and time-shifting properties. In particular, consider an LTI system for which the input and output satisfy a linear constant-coefficient difference equation of the form

$$\sum_{k=0}^{N} a_k y[n-k] = \sum_{k=0}^{M} b_k x[n-k]. \tag{10.105}$$

Then taking z-transforms of both sides of eq. (10.105) and using the linearity and time-shifting properties, we obtain

$$\sum_{k=0}^{N} a_k z^{-k} Y(z) = \sum_{k=0}^{M} b_k z^{-k} X(z),$$

or

$$Y(z) \sum_{k=0}^{N} a_k z^{-k} = X(z) \sum_{k=0}^{M} b_k z^{-k},$$

so that

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{k=0}^{M} b_k z^{-k}}{\sum_{k=0}^{N} a_k z^{-k}}. \tag{10.106}$$

We note in particular that the system function for a system satisfying a linear constant-coefficient difference equation is always rational. Consistent with our previous example and with the related discussion for the Laplace transform, the difference equation by itself does not provide information about which ROC to associate with the algebraic expression $H(z)$. An additional constraint, such as the causality or stability of the system, however, serves to specify the region of convergence. For example, if we know in addition that the system is causal, the ROC will be outside the outermost pole. If the system is stable, the ROC must include the unit circle.

10.7.4 Examples Relating System Behavior to the System Function

As the previous subsections illustrate, many properties of discrete-time LTI systems can be directly related to the system function and its characteristics. In this section, we give several additional examples to show how z-transform properties can be used in analyzing systems.

Example 10.26

Suppose that we are given the following information about an LTI system:

1. If the input to the system is $x_1[n] = (1/6)^n u[n]$, then the output is

$$y_1[n] = \left[a\left(\frac{1}{2}\right)^n + 10\left(\frac{1}{3}\right)^n\right] u[n],$$

where a is a real number.

2. If $x_2[n] = (-1)^n$, then the output is $y_2[n] = \frac{7}{4}(-1)^n$. As we now show, from these two pieces of information, we can determine the system function $H(z)$ for this system, including the value of the number a, and can also immediately deduce a number of other properties of the system.

The z-transforms of the signals specified in the first piece of information are

$$X_1(z) = \frac{1}{1-\frac{1}{6}z^{-1}}, \quad |z| > \frac{1}{6}, \tag{10.107}$$

$$\begin{aligned} Y_1(z) &= \frac{a}{1-\frac{1}{2}z^{-1}} + \frac{10}{1-\frac{1}{3}z^{-1}} \\ &= \frac{(a+10)-(5+\frac{a}{3})z^{-1}}{(1-\frac{1}{2}z^{-1})(1-\frac{1}{3}z^{-1})}, \quad |z| > \frac{1}{2}. \end{aligned} \tag{10.108}$$

From eq. (10.96), it follows that the algebraic expression for the system function is

$$H(z) = \frac{Y_1(z)}{X_1(z)} = \frac{[(a+10)-(5+\frac{a}{3})z^{-1}][1-\frac{1}{6}z^{-1}]}{(1-\frac{1}{2}z^{-1})(1-\frac{1}{3}z^{-1})}. \tag{10.109}$$

Furthermore, we know that the response to $x_2[n] = (-1)^n$ must equal $(-1)^n$ multiplied by the system function $H(z)$ evaluated at $z = -1$. Thus from the second piece of information given, we see that

$$\frac{7}{4} = H(-1) = \frac{[(a+10)+5+\frac{a}{3}][\frac{7}{6}]}{(\frac{3}{2})(\frac{4}{3})}. \tag{10.110}$$

Solving eq. (10.110), we find that $a = -9$, so that

$$H(z) = \frac{(1-2z^{-1})(1-\frac{1}{6}z^{-1})}{(1-\frac{1}{2}z^{-1})(1-\frac{1}{3}z^{-1})}, \tag{10.111}$$

or

$$H(z) = \frac{1-\frac{13}{6}z^{-1}+\frac{1}{3}z^{-2}}{1-\frac{5}{6}z^{-1}+\frac{1}{6}z^{-2}}, \tag{10.112}$$

or, finally,

$$H(z) = \frac{z^2-\frac{13}{6}z+\frac{1}{3}}{z^2-\frac{5}{6}z+\frac{1}{6}}. \tag{10.113}$$

Also, from the convolution property, we know that the ROC of $Y_1(z)$ must include at least the intersections of the ROCs of $X_1(z)$ and $H(z)$. Examining the three possible ROCs for $H(z)$ (namely, $|z| < 1/3$, $1/3 < |z| < 1/2$, and $|z| > 1/2$), we find that the only choice that is consistent with the ROCs of $X_1(z)$ and $Y_1(z)$ is $|z| > 1/2$.

Since the ROC for the system includes the unit circle, we know that the system is stable. Furthermore, from eq. (10.113) with $H(z)$ viewed as a ratio of polynomials in z, the order of the numerator does not exceed that of the denominator, and thus we can conclude that the LTI system is causal. Also, using eqs. (10.112) and (10.106), we can write the difference equation that, together with the condition of initial rest, characterizes the system:

$$y[n] - \frac{5}{6}y[n-1] + \frac{1}{6}y[n-2] = x[n] - \frac{13}{6}x[n-1] + \frac{1}{3}x[n-2].$$

Example 10.27

Consider a stable and causal system with impulse response $h[n]$ and rational system function $H(z)$. Suppose it is known that $H(z)$ contains a pole at $z = 1/2$ and a zero somewhere on the unit circle. The precise number and locations of all of the other poles and zeros are unknown. For each of the following statements, let us determine whether we can definitely say that it is true, whether we can definitely say that it is false, or whether there is insufficient information given to determine if it is true or not:

(a) $\mathfrak{F}\{(1/2)^n h[n]\}$ converges.
(b) $H(e^{j\omega}) = 0$ for some ω.
(c) $h[n]$ has finite duration.
(d) $h[n]$ is real.
(e) $g[n] = n[h[n] * h[n]]$ is the impulse response of a stable system.

Statement (a) is true. $\mathfrak{F}\{(1/2)^n h[n]\}$ corresponds to the value of the z-transform of $h[n]$ at $z = 2$. Thus, its convergence is equivalent to the point $z = 2$ being in the ROC. Since the system is stable and causal, all of the poles of $H(z)$ are inside the unit circle, and the ROC includes all the points outside the unit circle, including $z = 2$.

Statement (b) is true because there is a zero on the unit circle.

Statement (c) is false because a finite-duration sequence must have an ROC that includes the entire z-plane, except possibly $z = 0$ and/or $z = \infty$. This is not consistent with having a pole at $z = 1/2$.

Statement (d) requires that $H(z) = H^*(z^*)$. This in turn implies that if there is a pole (zero) at a nonreal location $z = z_0$, there must also be a pole (zero) at $z = z_0^*$. Insufficient information is given to validate such a conclusion.

Statement (e) is true. Since the system is causal, $h[n] = 0$ for $n < 0$. Consequently, $h[n] * h[n] = 0$ for $n < 0$; i.e., the system with $h[n] * h[n]$ as its impulse response is causal. The same is then true for $g[n] = n[h[n] * h[n]]$. Furthermore, by the convolution property set forth in Section 10.5.7, the system function corresponding to the impulse response $h[n] * h[n]$ is $H^2(z)$, and by the differentiation property presented in Section 10.5.8, the system function corresponding to $g[n]$ is

$$G(z) = -z\frac{d}{dz}H^2(z) = -2zH(z)\left[\frac{d}{dz}H(z)\right]. \tag{10.114}$$

From eq. (10.114), we can conclude that the poles of $G(z)$ are at the same locations as those of $H(z)$, with the possible exception of the origin. Therefore, since $H(z)$ has all its poles inside the unit circle, so must $G(z)$. It follows that $g[n]$ is the impulse response of a causal and stable system.

10.8 SYSTEM FUNCTION ALGEBRA AND BLOCK DIAGRAM REPRESENTATIONS

Just as with the Laplace transform in continuous time, the z-transform in discrete time allows us to replace time-domain operations such as convolution and time shifting with algebraic operations. This was exploited in Section 10.7.3, where we were able to replace the difference-equation description of an LTI system with an algebraic description. The use of the z-transform to convert system descriptions to algebraic equations is also helpful in analyzing interconnections of LTI systems and in representing and synthesizing systems as interconnections of basic system building blocks.

10.8.1 System Functions for Interconnections of LTI Systems

The system function algebra for analyzing discrete-time block diagrams such as series, parallel, and feedback interconnections is exactly the same as that for the corresponding continuous-time systems in Section 9.8.1. For example, the system function for the cascade of two discrete-time LTI systems is the product of the system functions for the individual systems in the cascade. Also, consider the feedback interconnection of two systems, as shown in Figure 10.17. It is relatively involved to determine the difference equation or impulse response for the overall system working directly in the time domain. However, with the systems and sequences expressed in terms of their z-transforms, the analysis involves only algebraic equations. The specific equations for the interconnection of Figure 10.17 exactly parallel eqs. (9.159)–(9.163), with the final result that the overall system function for the feedback system of Figure 10.17 is

$$\frac{Y(z)}{X(z)} = H(z) = \frac{H_1(z)}{1 + H_1(z)H_2(z)}. \tag{10.115}$$

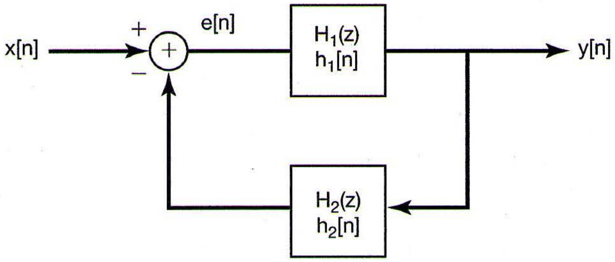


Figure 10.17 Feedback interconnection of two systems.

10.8.2 Block Diagram Representations for Causal LTI Systems Described by Difference Equations and Rational System Functions

As in Section 9.8.2, we can represent causal LTI systems described by difference equations using block diagrams involving three basic operations—in this case, addition, multiplication by a coefficient, and a unit delay. In Section 2.4.3, we described such a block diagram for a first-order difference equation. We first revisit that example, this time using system function algebra, and then consider several slightly more complex examples to illustrate the basic ideas in constructing block diagram representations.

Example 10.28

Consider the causal LTI system with system function

$$H(z) = \frac{1}{1 - \frac{1}{4}z^{-1}}. \tag{10.116}$$

Using the results in Section 10.7.3, we find that this system can also be described by the difference equation

$$y[n] - \frac{1}{4}y[n-1] = x[n],$$

together with the condition of initial rest. In Section 2.4.3 we constructed a block diagram representation for a first-order system of this form, and an equivalent block diagram (corresponding to Figure 2.28 with $a = -1/4$ and $b = 1$) is shown in Figure 10.18(a). Here, z^{-1} is the system function of a unit delay. That is, from the time-shifting property, the input and output of this system are related by

$$w[n] = y[n-1].$$

The block diagram in Figure 10.18(a) contains a feedback loop much as for the system considered in the previous subsection and pictured in Figure 10.17. In fact, with some minor modifications, we can obtain the equivalent block diagram shown in Figure 10.18(b), which is exactly in the form shown in Figure 10.17, with $H_1(z) = 1$ and $H_2(z) = -1/4z^{-1}$. Then, applying eq. (10.115), we can verify that the system function of the system in Figure 10.18 is given by eq. (10.116).

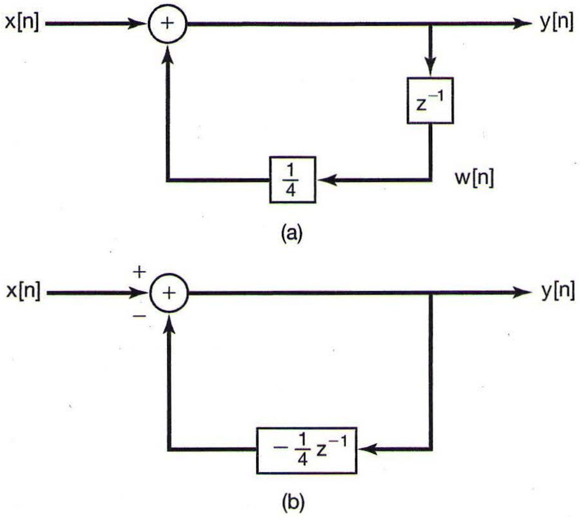


Figure 10.18 (a) Block diagram representations of the causal LTI system in Example 10.28; (b) equivalent block diagram representation.

Example 10.29

Suppose we now consider the causal LTI system with system function

$$H(z) = \frac{1-2z^{-1}}{1-\frac{1}{4}z^{-1}} = \left(\frac{1}{1-\frac{1}{4}z^{-1}}\right)(1-2z^{-1}). \tag{10.117}$$

As eq. (10.117) suggests, we can think of this system as the cascade of a system with system function $1/[1-(1/4)z^{-1}]$ and one with system function $1-2z^{-1}$. We have illustrated the cascade in Figure 10.19(a), in which we have used the block diagram in Figure 10.18(a) to represent $1/[1-(1/4)z^{-1}]$. We have also represented $1-2z^{-1}$ using a unit delay, an adder, and a coefficient multiplier. Using the time-shifting property, we then see that the input $v[n]$ and output $y[n]$ of the system with system function $1-2z^{-1}$

are related by

$$y[n] = v[n] - 2v[n-1].$$

While the block diagram in Figure 10.19(a) is certainly a valid representation of the system in eq. (10.117), it has an inefficiency whose elimination leads to an alternative block-diagram representation. To see this, note that the input to both unit delay elements in Figure 10.19(a) is $v[n]$, so that the outputs of these elements are identical; i.e.,

$$w[n] = s[n] = v[n-1].$$

Consequently, we need not keep both of these delay elements, and we can simply use the output of one of them as the signal to be fed to both coefficient multipliers. The result is the block diagram representation in Figure 10.19(b). Since each unit delay element requires a memory register to store the preceding value of its input, the representation in Figure 10.19(b) requires less memory than that in Figure 10.19(a).

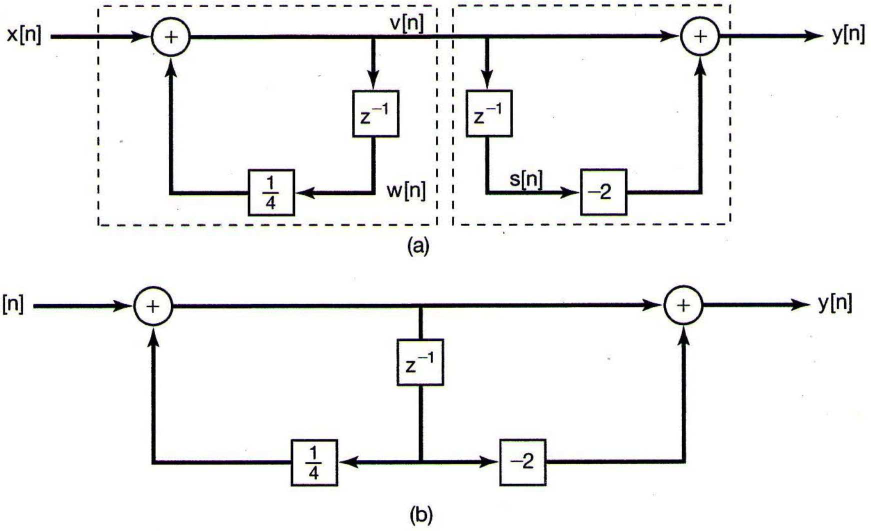


Figure 10.19 (a) Block-diagram representations for the system in Example 10.29; (b) equivalent block-diagram representation using only one unit delay element.

Example 10.30

Next, consider the second-order system function

$$H(z) = \frac{1}{(1+\frac{1}{2}z^{-1})(1-\frac{1}{4}z^{-1})} = \frac{1}{1+\frac{1}{4}z^{-1}-\frac{1}{8}z^{-2}}, \qquad (10.118)$$

which is also described by the difference equation

$$y[n] + \frac{1}{4}y[n-1] - \frac{1}{8}y[n-2] = x[n]. \qquad (10.119)$$

Using the same ideas as in Example 10.28, we obtain the block-diagram representation for this system shown in Figure 10.20(a). Specifically, since the two system function blocks in this figure with system function z^{-1} are unit delays, we have

$$\begin{aligned} f[n] &= y[n-1], \\ e[n] &= f[n-1] = y[n-2], \end{aligned}$$

so that eq. (10.119) can be rewritten as

$$y[n] = -\frac{1}{4}y[n-1] + \frac{1}{8}y[n-2] + x[n],$$

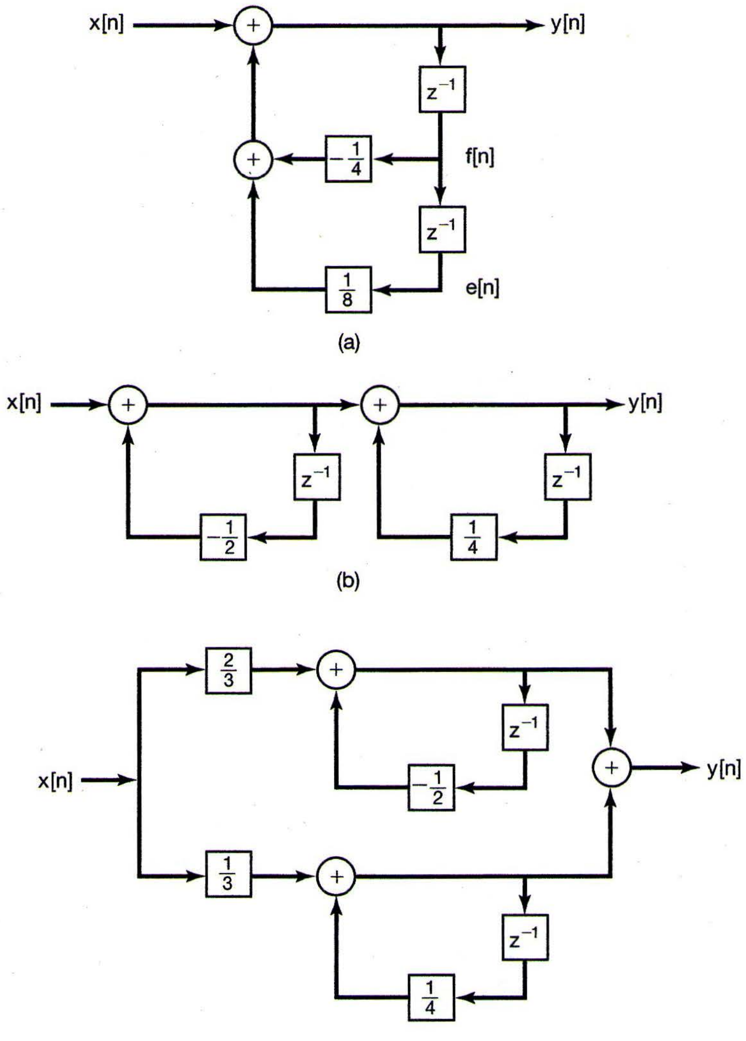

x[n]
y[n]
z^{-1}
$-\frac{1}{4}$
f[n]
z^{-1}
$\frac{1}{8}$
e[n]
(a)
x[n]
y[n]
z^{-1}
$-\frac{1}{2}$
z^{-1}
$\frac{1}{4}$
(b)
$\frac{2}{3}$
z^{-1}
$-\frac{1}{2}$
x[n]
y[n]
$\frac{1}{3}$
z^{-1}
$\frac{1}{4}$

Figure 10.20 Block-diagram representations for the system in Example 10.30: (a) direct form; (b) cascade form; (c) parallel form.

or

$$y[n] = -\frac{1}{4}f[n] + \frac{1}{8}e[n] + x[n],$$

which is exactly what the figure represents.

The block diagram in Figure 10.20(a) is commonly referred to as a *direct-form* representation, since the coefficients appearing in the diagram can be determined by inspection from the coefficients appearing in the difference equation or, equivalently, the system function. Alternatively, as in continuous time, we can obtain both *cascade-form* and *parallel-form* block diagrams with the aid of a bit of system function algebra. Specifically, we can rewrite eq. (10.118) as

$$H(z) = \left(\frac{1}{1 + \frac{1}{2}z^{-1}}\right)\left(\frac{1}{1 - \frac{1}{4}z^{-1}}\right), \tag{10.120}$$

which suggests the cascade-form representation depicted in Figure 10.20(b) in which the system is represented as the cascade of two systems corresponding to the two factors in eq. (10.120).

Also, by performing a partial-fraction expansion, we obtain

$$H(z) = \frac{\frac{2}{3}}{1 + \frac{1}{2}z^{-1}} + \frac{\frac{1}{3}}{1 - \frac{1}{4}z^{-1}},$$

which leads to the parallel-form representation depicted in Figure 10.20(c).

Example 10.31

Finally, consider the system function

$$H(z) = \frac{1 - \frac{7}{4}z^{-1} - \frac{1}{2}z^{-2}}{1 + \frac{1}{4}z^{-1} - \frac{1}{8}z^{-2}}. \tag{10.121}$$

Writing

$$H(z) = \left(\frac{1}{1 + \frac{1}{4}z^{-1} - \frac{1}{8}z^{-2}}\right)\left(1 - \frac{7}{4}z^{-1} - \frac{1}{2}z^{-2}\right) \tag{10.122}$$

suggests representing the system as the cascade of the system in Figure 10.20(a) and the system with system function $1 - \frac{7}{4}z^{-1} - \frac{1}{2}z^{-2}$. However, as in Example 10.29, the unit delay elements needed to implement the first term in eq. (10.122) also produce the delayed signals needed in computing the output of the second system. The result is the direct-form block diagram shown in Figure 10.21, the details of the construction of which are examined in Problem 10.38. The coefficients in the direct-form representation can be determined by inspection from the coefficients in the system function of eq. (10.121).

We can also write $H(z)$ in the forms

$$H(z) = \left(\frac{1 + \frac{1}{4}z^{-1}}{1 + \frac{1}{2}z^{-1}}\right)\left(\frac{1 - 2z^{-1}}{1 - \frac{1}{4}z^{-1}}\right) \tag{10.123}$$

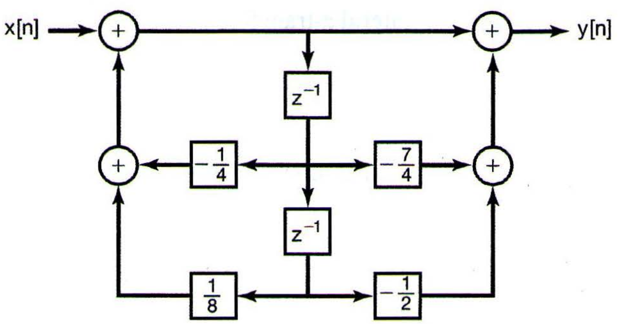


Figure 10.21 Direct-form representation for the system in Example 10.31.

and

$$H(z) = 4 + \frac{5/3}{1 + \frac{1}{2}z^{-1}} - \frac{14/3}{1 - \frac{1}{4}z^{-1}}. \qquad (10.124)$$

Eq. (10.123) suggests a cascade-form representation, while eq. (10.124) leads to a parallel-form block diagram. These are also considered in Problem 10.38.

The concepts used in constructing block-diagram representations in the preceding examples can be applied directly to higher order systems, and several examples are considered in Problem 10.39. As in continuous time, there is typically considerable flexibility in doing this—e.g., in how numerator and denominator factors are paired in a product representation as in eq. (10.123), in the way in which each factor is implemented, and in the order in which the factors are cascaded. While all of these variations lead to representations of the same system, in practice there are differences in the behavior of the different block diagrams. Specifically, each block-diagram representation of a system can be translated directly into a computer algorithm for the implementation of the system. However, because the finite word length of a computer necessitates quantizing the coefficients in the block diagram and because there is numerical roundoff as the algorithm operates, each of these representations will lead to an algorithm that only approximates the behavior of the original system. Moreover, the errors in each of these approximations will be somewhat different. Because of these differences, considerable effort has been put into examining the relative merits of the various block-diagram representations in terms of their accuracy and sensitivity to quantization effects. For discussions of this subject, the reader may turn to the references on digital signal processing in the bibliography at the end of the book.

10.9 THE UNILATERAL z-TRANSFORM

The form of the z-transform considered thus far in this chapter is often referred to as the *bilateral z-transform.* As was the case with the Laplace transform, there is an alternative form, referred to as the *unilateral z-transform,* that is particularly useful in analyzing causal systems specified by linear constant-coefficient difference equations with nonzero initial conditions (i.e., systems that are not initially at rest). In this section, we introduce the unilateral z-transform and illustrate some of its properties and uses, paralleling our discussion of the unilateral Laplace transform in Section 9.9.

The unilateral z-transform of a sequence $x[n]$ is defined as

$$\mathcal{X}(z) = \sum_{n=0}^{\infty} x[n]z^{-n}. \tag{10.125}$$

As in previous chapters, we adopt a convenient shorthand notation for a signal and its unilateral z-transform:

$$x[n] \overset{\mathcal{UZ}}{\longleftrightarrow} \mathcal{X}(z) = \mathcal{UZ}\{x[n]\}. \tag{10.126}$$

The unilateral z-transform differs from the bilateral transform in that the summation is carried out only over nonnegative values of n, whether or not $x[n]$ is zero for $n < 0$. Thus the unilateral z-transform of $x[n]$ can be thought of as the bilateral transform of $x[n]u[n]$ (i.e., $x[n]$ multiplied by a unit step). In particular, then, for any sequence that is zero for $n < 0$, the unilateral and bilateral z-transforms will be identical. Referring to the discussion of regions of convergence in Section 10.2, we also see that, since $x[n]u[n]$ is always a right-sided sequence, the region of convergence of $\mathcal{X}(z)$ is always the exterior of a circle.

Because of the close connection between bilateral and unilateral z-transforms, the calculation of unilateral transforms proceeds much as for bilateral transforms, with the caveat that we must take care to limit the range of summation in the transform to $n \geq 0$. Similarly, the calculation of inverse unilateral transforms is basically the same as for bilateral transforms, once we take into account the fact that the ROC for a unilateral transform is *always* the exterior of a circle.

10.9.1 Examples of Unilateral z-Transforms and Inverse Transforms

Example 10.32

Consider the signal

$$x[n] = a^n u[n]. \tag{10.127}$$

Since $x[n] = 0, n < 0$, the unilateral and bilateral transforms are equal for this example, and thus, in particular,

$$\mathcal{X}(z) = \frac{1}{1 - az^{-1}}, \qquad |z| > |a|. \tag{10.128}$$

Example 10.33

Let

$$x[n] = a^{n+1}u[n+1]. \tag{10.129}$$

In this case the unilateral and bilateral transforms are *not* equal, since $x[-1] = 1 \neq 0$. The bilateral transform is obtained from Example 10.1 and the time-shifting property set forth in Section 10.5.2. Specifically,

$$X(z) = \frac{z}{1 - az^{-1}}, \qquad |z| > |a|. \tag{10.130}$$

In contrast, the unilateral transform is

$$\mathcal{X}(z) = \sum_{n=0}^{\infty} x[n]z^{-n}$$

$$= \sum_{n=0}^{\infty} a^{n+1}z^{-n},$$

or

$$\mathcal{X}(z) = \frac{a}{1 - az^{-1}}, \qquad |z| > |a|. \tag{10.131}$$

Example 10.34

Consider the unilateral z-transform

$$\mathcal{X}(z) = \frac{3 - \frac{5}{6}z^{-1}}{(1 - \frac{1}{4}z^{-1})(1 - \frac{1}{3}z^{-1})}. \tag{10.132}$$

In Example 10.9, we considered the inverse transform for a bilateral z-transform $X(z)$ of the same form as in eq. (10.132) and for several different ROCs. In the case of the unilateral transform, the ROC must be the exterior of the circle of radius equal to the largest magnitude of the poles of $\mathcal{X}(z)$—in this instance, all points z with $|z| > 1/3$. We can then invert the unilateral transform exactly as in Example 10.9, yielding

$$x[n] = \left(\frac{1}{4}\right)^n u[n] + 2\left(\frac{1}{3}\right)^n u[n] \quad \text{for} \quad n \geq 0. \tag{10.133}$$

In eq. (10.133), we have emphasized the fact that inverse unilateral z-transforms provide us with information about $x[n]$ *only* for $n \geq 0$.

Another approach to inverse transforms introduced in Section 10.3, namely, identifying the inverse transforms from the coefficients in the power-series expansion of the z-transform, also can be used for unilateral transforms. However, in the unilateral case, a constraint which must be satisfied is that, as a consequence of eq. (10.125), the power-series expansion for the transform cannot contain terms with positive powers of z. For instance, in Example 10.13 we performed long division on the bilateral transform

$$X(z) = \frac{1}{1 - az^{-1}} \tag{10.134}$$

in two ways, corresponding to the two possible ROCs for $X(z)$. Only one of these choices, namely, that corresponding to the ROC $|z| > |a|$, led to a series expansion without positive powers of z, i.e.,

$$\frac{1}{1 - az^{-1}} = 1 + az^{-1} + a^2z^{-2} + \ldots, \tag{10.135}$$

and this is the *only* choice for the expansion if eq. (10.134) represents a unilateral transform.

Note that the requirement that $\mathcal{X}(z)$ have a power-series expansion with no terms with positive powers of z implies that not every function of z can *be* a unilateral z-transform. In particular, if we consider a rational function of z written as a ratio of polynomials in z (not in z^{-1}), i.e.,

$$\frac{p(z)}{q(z)}, \tag{10.136}$$

then for this to be a unilateral transform (with the appropriately chosen ROC as the exterior of a circle), the degree of the numerator must be no bigger than the degree of the denominator.

Example 10.35

A simple example illustrating the preceding point is given by the rational function in eq. (10.130), which we can write as a ratio of polynomials in z:

$$\frac{z^2}{z-a}. \tag{10.137}$$

There are two possible bilateral transforms that can be associated with this function, namely those corresponding to the two possible ROCs, $|z| < |a|$ and $|z| > |a|$. The choice $|z| > |a|$ corresponds to a right-sided sequence, but *not* to a signal that is zero for all $n < 0$, since its inverse transform, which is given by eq. (10.129), is nonzero for $n = -1$.

More generally, if we associate eq. (10.136) with the bilateral transform with the ROC that is the exterior of the circle with radius given by the magnitude of the largest root of $q(z)$, then the inverse transform will certainly be right sided. However, for it to be zero for *all* $n < 0$, it must also be the case that degree$(p(z)) \leq$ degree$(q(z))$.

10.9.2 Properties of the Unilateral z-Transform

The unilateral z-transform has many important properties, some of which are identical to their bilateral counterparts and several of which differ in significant ways. Table 10.3 summarizes these properties. Note that we have not included a column explicitly identifying the ROC for the unilateral z-transform for each signal, since the ROC of any unilateral z-transform is always the exterior of a circle. For example, the ROC for a rational unilateral z-transform is always outside the outermost pole.

By contrasting this table with the corresponding Table 10.1 for bilateral z-transforms, we can gain considerable insight into the nature of the unilateral transform. In particular, several properties—namely, linearity, scaling in the z-domain, time expansion, conjugation, and differentiation in the z-domain—are identical to their bilateral counterparts, as is the initial-value theorem stated in Section 10.5.9, which is fundamentally a unilateral transform property, since it requires $x[n] = 0$ for $n < 0$. One bilateral property, namely, the time-reversal property set forth in Section 10.5.4, obviously has no meaningful counterpart for the unilateral transform, while the remaining properties differ in important ways between the bilateral and unilateral cases.

TABLE 10.3 PROPERTIES OF THE UNILATERAL z-TRANSFORM

Property	Signal	Unilateral z-Transform
—	$x[n]$	$\mathcal{X}(z)$
—	$x_1[n]$	$\mathcal{X}_1(z)$
—	$x_2[n]$	$\mathcal{X}_2(z)$
Linearity	$ax_1[n] + bx_2[n]$	$a\mathcal{X}_1(z) + b\mathcal{X}_2(z)$
Time delay	$x[n-1]$	$z^{-1}\mathcal{X}(z) + x[-1]$
Time advance	$x[n+1]$	$z\mathcal{X}(z) - zx[0]$
Scaling in the z-domain	$e^{j\omega_0 n}x[n]$	$\mathcal{X}(e^{-j\omega_0}z)$
	$z_0^n x[n]$	$\mathcal{X}(z/z_0)$
	$a^n x[n]$	$\mathcal{X}(a^{-1}z)$
Time expansion	$x_k[n] = \begin{cases} x[m], & n = mk \\ 0, & n \neq mk \end{cases}$ for any m	$\mathcal{X}(z^k)$
Conjugation	$x^*[n]$	$\mathcal{X}^*(z^*)$
Convolution (assuming that $x_1[n]$ and $x_2[n]$ are identically zero for $n < 0$)	$x_1[n] * x_2[n]$	$\mathcal{X}_1(z)\mathcal{X}_2(z)$
First difference	$x[n] - x[n-1]$	$(1 - z^{-1})\mathcal{X}(z) - x[-1]$
Accumulation	$\sum_{k=0}^{n} x[k]$	$\frac{1}{1 - z^{-1}}\mathcal{X}(z)$
Differentiation in the z-domain	$nx[n]$	$-z\frac{d\mathcal{X}(z)}{dz}$

Initial Value Theorem

$$x[0] = \lim_{z\to\infty} \mathcal{X}(z)$$

Let us examine the difference in the convolution property first. Table 10.3 states that if $x_1[n] = x_2[n] = 0$ for all $n < 0$, then

$$x_1[n] * x_2[n] \overset{\mathcal{UZ}}{\longleftrightarrow} \mathcal{X}_1(z)\mathcal{X}_2(z). \tag{10.138}$$

Since in this case the unilateral and bilateral transforms are identical for each of these signals, eq. (10.138) follows from the bilateral convolution property. Thus, the system analysis and system function algebra developed and used in this chapter apply without change to unilateral transforms, as long as we are considering causal LTI systems (for which the system function is both the bilateral *and* the unilateral transform of the impulse response) with inputs that are identically zero for $n < 0$. An example of such application is to the accumulation or summation property in Table 10.3. Specifically, if $x[n] = 0$ for $n < 0$, then

$$\sum_{k=0}^{n} x[k] = x[n] * u[n] \overset{\mathcal{UZ}}{\longleftrightarrow} \mathcal{X}(z)\mathcal{U}(z) = \mathcal{X}(z)\frac{1}{1 - z^{-1}}. \tag{10.139}$$

As a second example, consider the following:

Example 10.36

Consider the causal LTI system described by the difference equation

$$y[n] + 3y[n-1] = x[n], \tag{10.140}$$

together with the condition of initial rest. The system function for this system is

$$\mathcal{H}(z) = \frac{1}{1+3z^{-1}}. \tag{10.141}$$

Suppose that the input to the system is $x[n] = \alpha u[n]$, where α is a given constant. In this case, the unilateral (and bilateral) z-transform of the output $y[n]$ is

$$\begin{aligned}\mathcal{Y}(z) = \mathcal{H}(z)\mathcal{X}(z) &= \frac{\alpha}{(1+3z^{-1})(1-z^{-1})} \\ &= \frac{(3/4)\alpha}{1+3z^{-1}} + \frac{(1/4)\alpha}{1-z^{-1}}.\end{aligned} \tag{10.142}$$

Applying Example 10.32 to each term of eq. (10.142) yields

$$y[n] = \alpha\left[\frac{1}{4} + \left(\frac{3}{4}\right)(-3)^n\right]u[n]. \tag{10.143}$$

An important point to note here is that the convolution property for unilateral z-transforms applies *only* if the signals $x_1[n]$ and $x_2[n]$ in eq. (10.138) are both identically zero for $n < 0$. While it is generally true that the bilateral transform of $x_1[n] * x_2[n]$ equals the product of the bilateral transforms of $x_1[n]$ and $x_2[n]$, the unilateral transform of $x_1[n] * x_2[n]$ in general does not equal the product of the unilateral transforms if $x_1[n]$ or $x_2[n]$ is nonzero for $n < 0$. This point is explored further in Problem 10.41.

Much of the importance of the unilateral z-transform lies in its application to analyzing causal systems and, in particular, systems characterized by linear constant-coefficient difference equations with possibly nonzero initial conditions. In Section 10.7 we saw how the bilateral transform—particularly the shifting property for bilateral z-transforms—could be used to analyze and compute solutions for LTI systems characterized by such difference equations, together with the assumption of initial rest. As we will now see, the shifting property for unilateral transforms, which differs from its bilateral counterpart, plays an analogous role for initialized systems.

To develop the shifting property for the unilateral transform, consider the signal

$$y[n] = x[n-1]. \tag{10.144}$$

Then

$$\begin{aligned}\mathcal{Y}(z) &= \sum_{n=0}^{\infty} x[n-1]z^{-n} \\ &= x[-1] + \sum_{n=1}^{\infty} x[n-1]z^{-n} \\ &= x[-1] + \sum_{n=0}^{\infty} x[n]z^{-(n+1)},\end{aligned}$$

or

$$\mathcal{Y}(z) = x[-1] + z^{-1}\sum_{n=0}^{\infty} x[n]z^{-n}, \tag{10.145}$$

so that

$$\mathcal{Y}(z) = x[-1] + z^{-1}\mathcal{X}(z). \tag{10.146}$$

By repeated application of eq. (10.146), the unilateral transform of

$$w[n] = y[n-1] = x[n-2] \tag{10.147}$$

is

$$\mathcal{W}(z) = x[-2] + x[-1]z^{-1} + z^{-2}\mathcal{X}(z). \tag{10.148}$$

Continuing this iterative procedure, we can also determine the unilateral transform of $x[n-m]$ for any positive value of m.

Eq. (10.146) is sometimes referred to as the time delay property, since $y[n]$ in eq. (10.144) is a delayed version of $x[n]$. There is also a time advance property for unilateral transforms that relates the transform of an advanced version of $x[n]$ to $\mathcal{X}(z)$. Specifically, as shown in Problem 10.60,

$$x[n+1] \overset{\mathcal{UZ}}{\longleftrightarrow} z\mathcal{X}(z) - zx[0]. \tag{10.149}$$

10.9.3 Solving Difference Equations Using the Unilateral z-Transform

The following example illustrates the use of unilateral z-transforms and the time delay property to solve linear constant-coefficient difference equations with nonzero initial conditions:

Example 10.37

Consider again the difference equation (10.140) with $x[n] = \alpha u[n]$ and with the initial condition

$$y[-1] = \beta. \tag{10.150}$$

Applying the unilateral transform to both sides of eq. (10.140) and using the linearity and time delay properties, we obtain

$$\mathcal{Y}(z) + 3\beta + 3z^{-1}\mathcal{Y}(z) = \frac{\alpha}{1-z^{-1}}. \tag{10.151}$$

Solving for $\mathcal{Y}(z)$ yields

$$\mathcal{Y}(z) = -\frac{3\beta}{1+3z^{-1}} + \frac{\alpha}{(1+3z^{-1})(1-z^{-1})}. \tag{10.152}$$

Referring to Example 10.36 and, in particular, eq. (10.142), we see that the second term on the right-hand side of eq. (10.152) equals the unilateral z-transform of the response of the system when the initial condition in eq. (10.150) is zero ($\beta = 0$). That is, this term represents the response of the causal LTI system described by eq. (10.140), together with the condition of initial rest. As in continuous-time, this response is frequently referred to as the zero-state response, i.e., the response when the initial condition or state is zero.

The first term on the right-hand side of eq. (10.152) is interpreted as the unilateral transform of the zero-input response—i.e., the response of the system when the input is zero ($\alpha = 0$). The zero-input response is a linear function of the value β of the initial condition. Moreover, eq. (10.152) illustrates the fact that the solution of a linear constant-coefficient difference equation with nonzero initial state is the superposition of the zero-state and zero-input responses. The zero-state response, obtained by setting the initial condition to zero, corresponds to the response of the causal LTI system defined by the difference equation and the condition of initial rest. The zero-input response is the response to the initial condition alone with the input set to zero. Problems 10.20 and 10.42 provide other examples illustrating the use of unilateral transforms to solve difference equations with nonzero initial conditions.

Finally, for any values of α and β, we can expand $\mathcal{Y}(z)$ in eq. (10.152) by the method of partial fractions and invert the result to obtain $y[n]$. For example, if $\alpha = 8$ and $\beta = 1$,

$$\mathcal{Y}(z) = \frac{3}{1+3z^{-1}} + \frac{2}{1-z^{-1}}, \tag{10.153}$$

and applying the unilateral transform pair in Example 10.32 to each term yields

$$y[n] = [3(-3)^n + 2]u[n], \quad \text{for } n \geq 0. \tag{10.154}$$

10.10 SUMMARY

In this chapter, we have developed the z-transform for discrete-time signals and systems. The discussion and development closely paralleled the corresponding treatment of the Laplace transform for continuous-time signals, but with some important differences. For example, in the complex s-plane the Laplace transform reduces to the Fourier transform on the imaginary axis, whereas in the complex z-plane the z-transform reduces to the Fourier transform on the unit circle. For the Laplace transform the ROC consists of a strip or half-plane (i.e., a strip extending to infinity in one direction), whereas for the z-transform the ROC is a ring, perhaps extending outward to infinity or inward to include the origin. As with the Laplace transform, time-domain characteristics such as the right-sided, left-sided, or two-sided nature of a sequence and the causality or stability of an LTI system can be associated with properties of the region of convergence. In particular, for rational z-transforms, these time-domain characteristics can be associated with the pole locations in relation to the region of convergence.

Because of the properties of z-transforms, LTI systems, including those described by linear constant-coefficient difference equations, can be analyzed in the transform domain by algebraic manipulations. System function algebra also is a very useful tool for the analysis of interconnections of LTI systems and for the construction of block diagram representations of LTI systems described by difference equations.

For the most part in this chapter, we have focused on bilateral z-transforms. However, as with the Laplace transform, we have also introduced a second form of the z-transform known as the unilateral z-transform. The unilateral transform, which can be viewed as the bilateral transform of a signal whose values for $n < 0$ have been set to zero, is particularly useful for analyzing systems described by linear constant-coefficient difference equations with nonzero initial conditions.

Chapter 10 Problems

The first section of problems belongs to the basic category, and the answers are provided in the back of the book. The remaining three sections contain problems belonging to the basic, advanced, and extension categories, respectively.

BASIC PROBLEMS WITH ANSWERS

10.1. Determine the constraint on $r = |z|$ for each of the following sums to converge:

(a) $\sum_{n=-1}^{\infty} (\frac{1}{2})^{n+1} z^{-n}$ **(b)** $\sum_{n=1}^{\infty} (\frac{1}{2})^{-n+1} z^{n}$

(c) $\sum_{n=0}^{\infty} \{\frac{1+(-1)^n}{2}\} z^{-n}$ **(d)** $\sum_{n=-\infty}^{\infty} (\frac{1}{2})^{|n|} \cos(\frac{\pi}{4} n) z^{-n}$

10.2. Consider the signal

$$x[n] = \left(\frac{1}{5}\right)^n u[n-3].$$

Use eq. (10.3) to evaluate the z-transform of this signal, and specify the corresponding region of convergence.

10.3. Let

$$x[n] = (-1)^n u[n] + \alpha^n u[-n-n_0].$$

Determine the constraints on the complex number α and the integer n_0, given that the ROC of $X(z)$ is

$$1 < |z| < 2.$$

10.4. Consider the signal

$$x[n] = \begin{cases} (\frac{1}{3})^n \cos(\frac{\pi}{4} n), & n \leq 0 \\ 0, & n > 0 \end{cases}.$$

Determine the poles and ROC for $X(z)$.

10.5. For each of the following algebraic expressions for the z-transform of a signal, determine the number of zeros in the finite z-plane and the number of zeros at infinity.

(a) $\dfrac{z^{-1}(1-\frac{1}{2}z^{-1})}{(1-\frac{1}{3}z^{-1})(1-\frac{1}{4}z^{-1})}$

(b) $\dfrac{(1-z^{-1})(1-2z^{-1})}{(1-3z^{-1})(1-4z^{-1})}$

(c) $\dfrac{z^{-2}(1-z^{-1})}{(1-\frac{1}{4}z^{-1})(1+\frac{1}{4}z^{-1})}$

10.6. Let $x[n]$ be an absolutely summable signal with rational z-transform $X(z)$. If $X(z)$ is known to have a pole at $z = 1/2$, could $x[n]$ be

(a) a finite-duration signal?
(b) a left-sided signal?
(c) a right-sided signal?
(d) a two-sided signal?

10.7. Suppose that the algebraic expression for the z-transform of $x[n]$ is

$$X(z) = \frac{1-\frac{1}{4}z^{-2}}{(1+\frac{1}{4}z^{-2})(1+\frac{5}{4}z^{-1}+\frac{3}{8}z^{-2})}.$$

How many different regions of convergence could correspond to $X(z)$?

10.8. Let $x[n]$ be a signal whose rational z-transform $X(z)$ contains a pole at $z = 1/2$. Given that

$$x_1[n] = \left(\frac{1}{4}\right)^n x[n]$$

is absolutely summable and

$$x_2[n] = \left(\frac{1}{8}\right)^n x[n]$$

is not absolutely summable, determine whether $x[n]$ is left sided, right sided, or two sided.

10.9. Using partial-fraction expansion and the fact that

$$a^n u[n] \overset{z}{\longleftrightarrow} \frac{1}{1-az^{-1}}, \quad |z| > |a|,$$

find the inverse z-transform of

$$X(z) = \frac{1-\frac{1}{3}z^{-1}}{(1-z^{-1})(1+2z^{-1})}, \quad |z| > 2.$$

10.10. Consider the following algebraic expression for the z-transform $X(z)$ of a signal $x[n]$:

$$X(z) = \frac{1+z^{-1}}{1+\frac{1}{3}z^{-1}}.$$

(a) Assuming the ROC to be $|z| > 1/3$, use long division to determine the values of $x[0]$, $x[1]$, and $x[2]$.

(b) Assuming the ROC to be $|z| < 1/3$, use long division to determine the values of $x[0]$, $x[-1]$, and $x[-2]$.

10.11. Find the inverse z-transform of

$$X(z) = \frac{1}{1{,}024}\left[\frac{1{,}024 - z^{-10}}{1 - \frac{1}{2}z^{-1}}\right], \quad |z| > 0.$$

10.12. By considering the geometric interpretation of the magnitude of the Fourier transform from the pole-zero plot, determine, for each of the following z-transforms, whether the corresponding signal has an approximately lowpass, bandpass, or highpass characteristic:

(a) $X(z) = \dfrac{z^{-1}}{1 + \frac{8}{9}z^{-1}}, \quad |z| > \frac{8}{9}$

(b) $X(z) = \dfrac{1 + \frac{8}{9}z^{-1}}{1 - \frac{16}{9}z^{-1} + \frac{64}{81}z^{-2}}, \quad |z| > \frac{8}{9}$

(c) $X(z) = \dfrac{1}{1 + \frac{64}{81}z^{-2}}, \quad |z| > \frac{8}{9}$

10.13. Consider the rectangular signal

$$x[n] = \begin{cases} 1, & 0 \le n \le 5 \\ 0, & \text{otherwise} \end{cases}.$$

Let

$$g[n] = x[n] - x[n-1].$$

(a) Find the signal $g[n]$ and directly evaluate its z-transform.

(b) Noting that

$$x[n] = \sum_{k=-\infty}^{n} g[k],$$

use Table 10.1 to determine the z-transform of $x[n]$.

10.14. Consider the triangular signal

$$g[n] = \begin{cases} n-1, & 2 \le n \le 7 \\ 13-n, & 8 \le n \le 12 \\ 0, & \text{otherwise} \end{cases}.$$

(a) Determine the value of n_0 such that

$$g[n] = x[n] * x[n-n_0],$$

where $x[n]$ is the rectangular signal considered in Problem 10.13.

(b) Use the convolution and shift properties in conjunction with $X(z)$ found in Problem 10.13 to determine $G(z)$. Verify that your answer satisfies the initial-value theorem.

10.15. Let

$$y[n] = \left(\frac{1}{9}\right)^n u[n].$$

Determine two distinct signals such that each has a z-transform $X(z)$ which satisfies both of the following conditions:

1. $[X(z) + X(-z)]/2 = Y(z^2)$.
2. $X(z)$ has only one pole and only one zero in the z-plane.

10.16. Consider the following system functions for stable LTI systems. Without utilizing the inverse z-transform, determine in each case whether or not the corresponding system is causal.

(a) $\dfrac{1 - \frac{4}{3}z^{-1} + \frac{1}{2}z^{-2}}{z^{-1}(1 - \frac{1}{2}z^{-1})(1 - \frac{1}{3}z^{-1})}$

(b) $\dfrac{z - \frac{1}{2}}{z^2 + \frac{1}{2}z - \frac{3}{16}}$

(c) $\dfrac{z + 1}{z + \frac{4}{3} - \frac{1}{2}z^{-2} - \frac{2}{3}z^{-3}}$

10.17. Suppose we are given the following five facts about a particular LTI system S with impulse response $h[n]$ and z-transform $H(z)$:

1. $h[n]$ is real.
2. $h[n]$ is right sided.
3. $\lim_{z \to \infty} H(z) = 1$.
4. $H(z)$ has two zeros.
5. $H(z)$ has one of its poles at a nonreal location on the circle defined by $|z| = 3/4$.

Answer the following two questions:

(a) Is S causal? **(b)** Is S stable?

10.18. Consider a causal LTI system whose input $x[n]$ and output $y[n]$ are related through the block diagram representation shown in Figure P10.18.

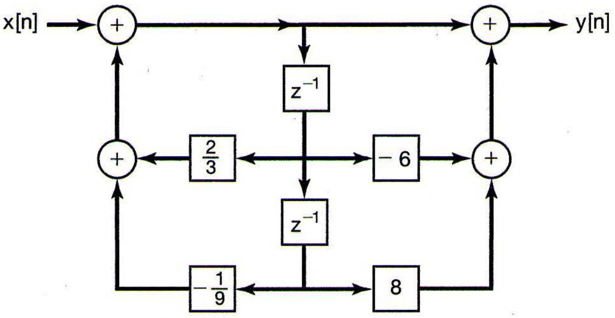


Figure P10.18

(a) Determine a difference equation relating $y[n]$ and $x[n]$.

(b) Is this system stable?

10.19. Determine the unilateral z-transform of each of the following signals, and specify the corresponding regions of convergence:

(a) $x_1[n] = (\frac{1}{4})^n u[n+5]$
(b) $x_2[n] = \delta[n+3] + \delta[n] + 2^n u[-n]$
(c) $x_3[n] = (\frac{1}{2})^{|n|}$

10.20. Consider a system whose input $x[n]$ and output $y[n]$ are related by

$$y[n-1] + 2y[n] = x[n].$$

(a) Determine the zero-input response of this system if $y[-1] = 2$.
(b) Determine the zero-state response of the system to the input $x[n] = (1/4)^n u[n]$.
(c) Determine the output of the system for $n \geq 0$ when $x[n] = (1/4)^n u[n]$ and $y[-1] = 2$.

BASIC PROBLEMS

10.21. Determine the z-transform for each of the following sequences. Sketch the pole-zero plot and indicate the region of convergence. Indicate whether or not the Fourier transform of the sequence exists.

(a) $\delta[n+5]$
(b) $\delta[n-5]$
(c) $(-1)^n u[n]$
(d) $(\frac{1}{2})^{n+1} u[n+3]$
(e) $(-\frac{1}{3})^n u[-n-2]$
(f) $(\frac{1}{4})^n u[3-n]$
(g) $2^n u[-n] + (\frac{1}{4})^n u[n-1]$
(h) $(\frac{1}{3})^{n-2} u[n-2]$

10.22. Determine the z-transform for the following sequences. Express all sums in closed form. Sketch the pole-zero plot and indicate the region of convergence. Indicate whether the Fourier transform of the sequence exists.

(a) $(\frac{1}{2})^n \{u[n+4] - u[n-5]\}$
(b) $n(\frac{1}{2})^{|n|}$
(c) $|n|(\frac{1}{2})^{|n|}$
(d) $4^n \cos[\frac{2\pi}{6}n + \frac{\pi}{4}]u[-n-1]$

10.23. Following are several z-transforms. For each one, determine the inverse z-transform using both the method based on the partial-fraction expansion and the Taylor's series method based on the use of long division.

$$X(z) = \frac{1 - z^{-1}}{1 - \frac{1}{4}z^{-2}}, \quad |z| > \frac{1}{2}.$$

$$X(z) = \frac{1 - z^{-1}}{1 - \frac{1}{4}z^{-2}}, \quad |z| < \frac{1}{2}.$$

$$X(z) = \frac{z^{-1} - \frac{1}{2}}{1 - \frac{1}{2}z^{-1}}, \quad |z| > \frac{1}{2}.$$

$$X(z) = \frac{z^{-1} - \frac{1}{2}}{1 - \frac{1}{2}z^{-1}}, \quad |z| < \frac{1}{2}.$$

$$X(z) = \frac{z^{-1} - \frac{1}{2}}{(1 - \frac{1}{2}z^{-1})^2}, \quad |z| > \frac{1}{2}.$$

$$X(z) = \frac{z^{-1} - \frac{1}{2}}{(1 - \frac{1}{2}z^{-1})^2}, \quad |z| < \frac{1}{2}.$$

10.24. Using the method indicated, determine the sequence that goes with each of the following z-transforms:

(a) Partial fractions:

$$X(z) = \frac{1 - 2z^{-1}}{1 + \frac{5}{2}z^{-1} + z^{-2}}, \text{ and } x[n] \text{ is absolutely summable.}$$

(b) Long division:

$$X(z) = \frac{1 - \frac{1}{2}z^{-1}}{1 + \frac{1}{2}z^{-1}}, \text{ and } x[n] \text{ is right sided.}$$

(c) Partial fractions:

$$X(z) = \frac{3}{z - \frac{1}{4} - \frac{1}{8}z^{-1}}, \text{ and } x[n] \text{ is absolutely summable.}$$

10.25. Consider a right-sided sequence $x[n]$ with z-transform

$$X(z) = \frac{1}{(1 - \frac{1}{2}z^{-1})(1 - z^{-1})}. \quad \text{(P10.25–1)}$$

(a) Carry out a partial-fraction expansion of eq. (P10.25–1) expressed as a ratio of polynomials in z^{-1}, and from this expansion, determine $x[n]$.

(b) Rewrite eq. (P10.25–1) as a ratio of polynomials in z, and carry out a partial-fraction expansion of $X(z)$ expressed in terms of polynomials in z. From this expansion, determine $x[n]$, and demonstrate that the sequence obtained is identical to that obtained in part (a).

10.26. Consider a left-sided sequence $x[n]$ with z-transform

$$X(z) = \frac{1}{(1 - \frac{1}{2}z^{-1})(1 - z^{-1})}.$$

(a) Write $X(z)$ as a ratio of polynomials in z instead of z^{-1}.

(b) Using a partial-fraction expression, express $X(z)$ as a sum of terms, where each term represents a pole from your answer in part (a).

(c) Determine $x[n]$.

10.27. A right-sided sequence $x[n]$ has z-transform

$$X(z) = \frac{3z^{-10} + z^{-7} - 5z^{-2} + 4z^{-1} + 1}{z^{-10} - 5z^{-7} + z^{-3}}.$$

Determine $x[n]$ for $n < 0$.

10.28. **(a)** Determine the z-transform of the sequence

$$x[n] = \delta[n] - 0.95\,\delta[n - 6].$$

(b) Sketch the pole-zero pattern for the sequence in part (a).

(c) By considering the behavior of the pole and zero vectors as the unit circle is traversed, develop an approximate sketch of the magnitude of the Fourier transform of $x[n]$.

10.29. By considering the geometric determination of the frequency response as discussed in Section 10.4, sketch, for each of the pole-zero plots in Figure P10.29, the magnitude of the associated Fourier transform.

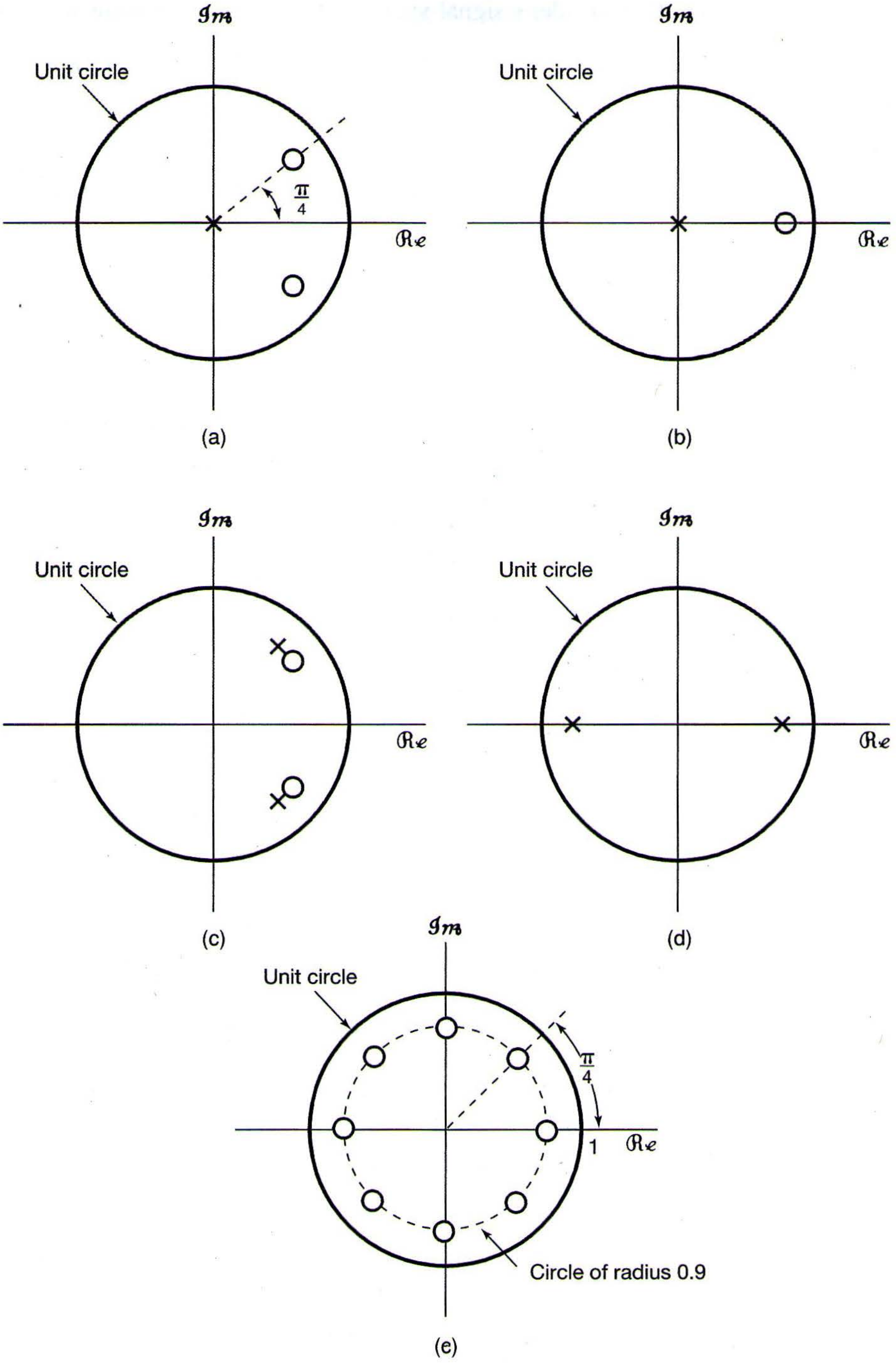

$\mathcal{Im}$
Unit circle
$\frac{\pi}{4}$
$\mathcal{Re}$
(a)
$\mathcal{Im}$
Unit circle
$\mathcal{Re}$
(b)
$\mathcal{Im}$
Unit circle
$\mathcal{Re}$
(c)
$\mathcal{Im}$
Unit circle
$\mathcal{Re}$
(d)
$\mathcal{Im}$
Unit circle
$\frac{\pi}{4}$
1
$\mathcal{Re}$
Circle of radius 0.9
(e)

Figure P10.29

10.30. Consider a signal $y[n]$ which is related to two signals $x_1[n]$ and $x_2[n]$ by

$$y[n] = x_1[n+3] * x_2[-n+1]$$

where

$$x_1[n] = \left(\frac{1}{2}\right)^n u[n] \quad \text{and} \quad x_2[n] = \left(\frac{1}{3}\right)^n u[n].$$

Given that

$$a^n u[n] \overset{z}{\longleftrightarrow} \frac{1}{1-az^{-1}}, \quad |z| > |a|,$$

use properties of the z-transform to determine the z-transform $Y(z)$ of $y[n]$.

10.31. We are given the following five facts about a discrete-time signal $x[n]$ with z-transform $X(z)$:

1. $x[n]$ is real and right-sided.
2. $X(z)$ has exactly two poles.
3. $X(z)$ has two zeros at the origin.
4. $X(z)$ has a pole at $z = \frac{1}{2}e^{j\pi/3}$.
5. $X(1) = \frac{8}{3}$.

Determine $X(z)$ and specify its region of convergence.

10.32. Consider an LTI system with impulse response

$$h[n] = \begin{cases} a^n, & n \geq 0 \\ 0, & n < 0 \end{cases}$$

and input

$$x[n] = \begin{cases} 1, & 0 \leq n \leq N-1 \\ 0, & \text{otherwise} \end{cases}.$$

(a) Determine the output $y[n]$ by explicitly evaluating the discrete convolution of $x[n]$ and $h[n]$.

(b) Determine the output $y[n]$ by computing the inverse z-transform of the product of the z-transforms of the input and the unit sample response.

10.33. **(a)** Determine the system function for the causal LTI system with difference equation

$$y[n] - \frac{1}{2}y[n-1] + \frac{1}{4}y[n-2] = x[n].$$

(b) Using z-transforms, determine $y[n]$ if

$$x[n] = \left(\frac{1}{2}\right)^n u[n].$$

10.34. A causal LTI system is described by the difference equation

$$y[n] = y[n-1] + y[n-2] + x[n-1].$$

(a) Find the system function $H(z) = Y(z)/X(z)$ for this system. Plot the poles and zeros of $H(z)$ and indicate the region of convergence.

(b) Find the unit sample response of the system.

(c) You should have found the system to be unstable. Find a stable (noncausal) unit sample response that satisfies the difference equation.

10.35. Consider an LTI system with input $x[n]$ and output $y[n]$ for which

$$y[n-1] - \frac{5}{2}y[n] + y[n+1] = x[n].$$

The system may or may not be stable or causal.

By considering the pole-zero pattern associated with the preceding difference equation, determine three possible choices for the unit sample response of the system. Show that each choice satisfies the difference equation.

10.36. Consider the linear, discrete-time, shift-invariant system with input $x[n]$ and output $y[n]$ for which

$$y[n-1] - \frac{10}{3}y[n] + y[n+1] = x[n].$$

The system is stable. Determine the unit sample response.

10.37. The input $x[n]$ and output $y[n]$ of a causal LTI system are related through the block-diagram representation shown in Figure P10.37.

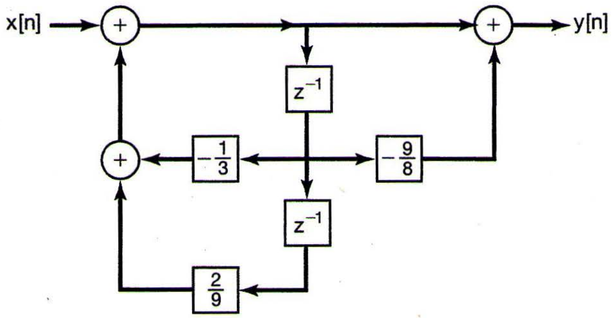


Figure P10.37

(a) Determine a difference equation relating $y[n]$ and $x[n]$.

(b) Is this system stable?

10.38. Consider a causal LTI system S with input $x[n]$ and a system function specified as

$$H(z) = H_1(z)H_2(z),$$

where

$$H_1(z) = \frac{1}{1 + \frac{1}{4}z^{-1} - \frac{1}{8}z^{-2}}$$

and

$$H_2(z) = 1 - \frac{7}{4}z^{-1} - \frac{1}{2}z^{-2}.$$

A block diagram corresponding to $H(z)$ may be obtained as a cascade connection of a block diagram for $H_1(z)$ followed by a block diagram for $H_2(z)$. The result is shown in Figure P10.38, in which we have also labeled the intermediate signals $e_1[n]$, $e_2[n]$, $f_1[n]$, and $f_2[n]$.

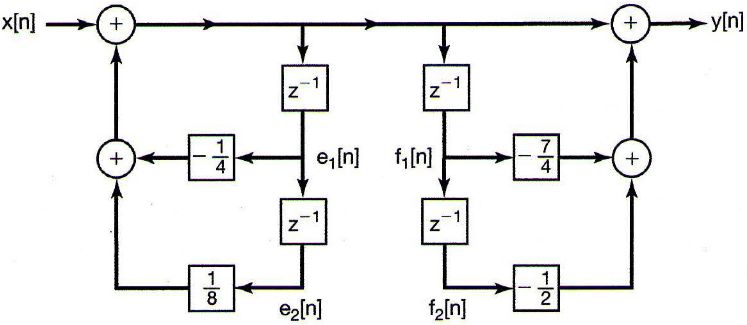


Figure P10.38

(a) How is $e_1[n]$ related to $f_1[n]$?
(b) How is $e_2[n]$ related to $f_2[n]$?
(c) Using your answers to the previous two parts as a guide, construct a direct-form block diagram for S that contains only two delay elements.
(d) Draw a cascade-form block diagram representation for S based on the observation that

$$H(z) = \left(\frac{1 + \frac{1}{4}z^{-1}}{1 + \frac{1}{2}z^{-1}}\right)\left(\frac{1 - 2z^{-1}}{1 - \frac{1}{4}z^{-1}}\right).$$

(e) Draw a parallel-form block diagram representation for S based on the observation that

$$H(z) = 4 + \frac{5/3}{1 + \frac{1}{2}z^{-1}} - \frac{14/3}{1 - \frac{1}{4}z^{-1}}.$$

10.39. Consider the following three system functions corresponding to causal LTI systems:

$$H_1(z) = \frac{1}{(1 - z^{-1} + \frac{1}{4}z^{-2})(1 - \frac{2}{3}z^{-1} + \frac{1}{9}z^{-2})},$$

$$H_2(z) = \frac{1}{(1 - z^{-1} + \frac{1}{2}z^{-2})(1 - \frac{1}{2}z^{-1} + z^{-2})},$$

$$H_3(z) = \frac{1}{(1 - z^{-1} + \frac{1}{2}z^{-2})(1 - z^{-1} + \frac{1}{4}z^{-2})}.$$

(a) For each system function, draw a direct-form block diagram.

(b) For each system function, draw a block diagram that corresponds to the cascade connection of two second-order block diagrams. Each second-order block diagram should be in direct form.

(c) For each system function, determine whether there exists a block diagram representation which is the cascade of four first-order block diagrams with the constraint that all the coefficient multipliers must be real.

10.40. Determine the unilateral z-transform for each of the sequences in Problem 10.21.

10.41. Consider the following two signals:

$$x_1[n] = \left(\frac{1}{2}\right)^{n+1} u[n+1],$$

$$x_2[n] = \left(\frac{1}{4}\right)^{n} u[n].$$

Let $\mathcal{X}_1(z)$ and $X_1(z)$ respectively be the unilateral and bilateral z-transforms of $x_1[n]$, and let $\mathcal{X}_2(z)$ and $X_2(z)$ respectively be the unilateral and bilateral z-transforms of $x_2[n]$.

(a) Take the inverse bilateral z-transform of $X_1(z)X_2(z)$ to determine $g[n] = x_1[n] * x_2[n]$.

(b) Take the inverse unilateral z-transform of $\mathcal{X}_1(z)\mathcal{X}_2(z)$ to obtain a signal $q[n]$ for $n \geq 0$. Observe that $q[n]$ and $g[n]$ are not identical for $n \geq 0$.

10.42. For each of the following difference equations and associated input and initial conditions, determine the zero-input and zero-state responses by using the unilateral z-transform:

(a) $y[n] + 3y[n-1] = x[n],$

$$x[n] = \left(\frac{1}{2}\right)^{n} u[n],$$

$$y[-1] = 1.$$

(b) $y[n] - \frac{1}{2}y[n-1] = x[n] - \frac{1}{2}x[n-1],$

$$x[n] = u[n],$$
$$y[-1] = 0.$$

(c) $y[n] - \frac{1}{2}y[n-1] = x[n] - \frac{1}{2}x[n-1],$

$$x[n] = u[n],$$
$$y[-1] = 1.$$

ADVANCED PROBLEMS

10.43. Consider an even sequence $x[n]$ (i.e., $x[n] = x[-n]$) with rational z-transform $X(z)$.

(a) From the definition of the z-transform, show that

$$X(z) = X\left(\frac{1}{z}\right).$$

(b) From your results in part (a), show that if a pole (zero) of $X(z)$ occurs at $z = z_0$, then a pole (zero) must also occur at $z = 1/z_0$.

(c) Verify the result in part (b) for each of the following sequences:

(1) $\delta[n+1] + \delta[n-1]$

(2) $\delta[n+1] - \frac{5}{2}\delta[n] + \delta[n-1]$

10.44. Let $x[n]$ be a discrete-time signal with z-transform $X(z)$. For each of the following signals, determine the z-transform in terms of $X(z)$:

(a) $\Delta x[n]$, where Δ is the first-difference operator defined by

$$\Delta x[n] = x[n] - x[n-1]$$

(b) $x_1[n] = \begin{cases} x[n/2], & n \text{ even} \\ 0, & n \text{ odd} \end{cases}$

(c) $x_1[n] = x[2n]$

10.45. Determine which of the following z-transforms could be the transfer function of a discrete-time linear system that is not necessarily stable, but for which the unit sample response is zero for $n < 0$. State your reasons clearly.

(a) $\dfrac{(1-z^{-1})^2}{1-\frac{1}{2}z^{-1}}$ **(b)** $\dfrac{(z-1)^2}{z-\frac{1}{2}}$

(c) $\dfrac{(z-\frac{1}{4})^5}{(z-\frac{1}{2})^6}$ **(d)** $\dfrac{(z-\frac{1}{4})^6}{(z-\frac{1}{2})^5}$

10.46. A sequence $x[n]$ is the output of an LTI system whose input is $s[n]$. The system is described by the difference equation

$$x[n] = s[n] - e^{8\alpha}s[n-8],$$

where $0 < \alpha < 1$.

(a) Find the system function

$$H_1(z) = \frac{X(z)}{S(z)},$$

and plot its poles and zeros in the z-plane. Indicate the region of convergence.

(b) We wish to recover $s[n]$ from $x[n]$ with an LTI system. Find the system function

$$H_2(z) = \frac{Y(z)}{X(z)}$$

such that $y[n] = s[n]$. Find all possible regions of convergence for $H_2(z)$, and for each, tell whether or not the system is causal or stable.

(c) Find all possible choices for the unit impulse response $h_2[n]$ such that

$$y[n] = h_2[n] * x[n] = s[n].$$

10.47. The following is known about a discrete-time LTI system with input $x[n]$ and output $y[n]$:

1. If $x[n] = (-2)^n$ for all n, then $y[n] = 0$ for all n.

2. If $x[n] = (1/2)^n u[n]$ for all n, then $y[n]$ for all n is of the form

$$y[n] = \delta[n] + a\left(\frac{1}{4}\right)^n u[n],$$

where a is a constant.

(a) Determine the value of the constant a.

(b) Determine the response $y[n]$ if the input $x[n]$ is

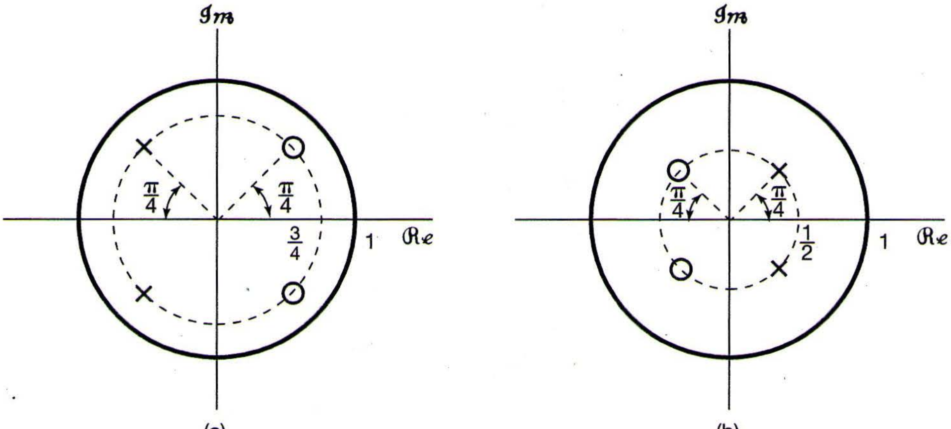
$$x[n] = 1, \text{ for all } n.$$

10.48. Suppose a second-order causal LTI system has been designed with a real impulse response $h_1[n]$ and a rational system function $H_1(z)$. The pole-zero plot for $H_1(z)$ is shown in Figure P10.48(a). Now consider another causal second-order system with impulse response $h_2[n]$ and rational system function $H_2(z)$. The pole-zero plot for $H_2(z)$ is shown in Figure P10.48(b). Determine a sequence $g[n]$ such that the following three conditions hold:

(1) $h_2[n] = g[n]h_1[n]$

(2) $g[n] = 0$ for $n < 0$

(3) $\sum_{k=0}^{\infty} |g[k]| = 3$



(a) (b)

Figure P10.48

10.49. In Property 4 of Section 10.2, it was stated that if $x[n]$ is a right-sided sequence and if the circle $|z| = r_0$ is in the ROC, then all finite values of z for which $|z| > r_0$ will also be in the ROC. In this discussion an intuitive explanation was given. A more formal argument parallels closely that used for Property 4 of Section 9.2, relating to the Laplace transform. Specifically, consider a right-sided sequence

$$x[n] = 0, \ n < N_1,$$

and for which

$$\sum_{n=-\infty}^{\infty} |x[n]| r_0^{-n} = \sum_{n=N_1}^{\infty} |x[n]| r_0^{-n} < \infty.$$

Then if $r_0 \leq r_1$,

$$\sum_{n=N_1}^{\infty} |x[n]| r_1^{-n} \leq A \sum_{n=N_1}^{\infty} |x[n]| r_0^{-n}, \qquad \text{(P10.49–1)}$$

where A is a positive constant.

(a) Show that eq. (P10.49–1) is true, and determine the constant A in terms of r_0, r_1, and N_1.

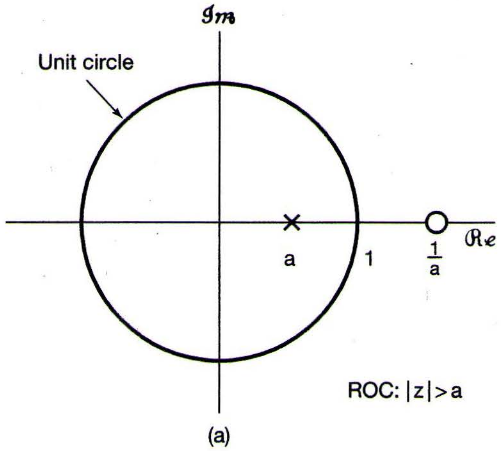
(b) From your result in part (a), show that Property 4 of Section 10.2 follows.

(c) Develop an argument similar to the foregoing one to demonstrate the validity of Property 5 of Section 10.2.

10.50. A discrete-time system with the pole-zero pattern shown in Figure P10.50(a) is referred to as a first-order all-pass system, since the magnitude of the frequency response is constant regardless of frequency.

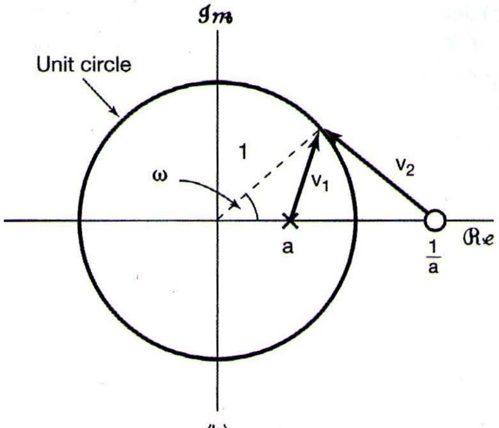
(a) Demonstrate algebraically that $|H(e^{j\omega})|$ is constant.

To demonstrate the same property geometrically, consider the vector diagram in Figure P10.50(b). We wish to show that the length of v_2 is proportional to the length of v_1 independently of the frequency ω.



(a)

Figure P10.50a



(b)

Figure P10.50b

(b) Express the length of v_1 using the law of cosines and the fact that v_1 is one leg of a triangle for which the other two legs are the unit vector and a vector of length a.

(c) In a manner similar to that in part (b), determine the length of v_2 and show that it is proportional in length to v_1 independently of ω.

10.51. Consider a real-valued sequence $x[n]$ with rational z-transform $X(z)$.

(a) From the definition of the z-transform, show that

$$X(z) = X^*(z^*).$$

(b) From your result in part (a), show that if a pole (zero) of $X(z)$ occurs at $z = z_0$, then a pole (zero) must also occur at $z = z_0^*$.

(c) Verify the result in part (b) for each of the following sequences:

(1) $x[n] = (\frac{1}{2})^n u[n]$

(2) $x[n] = \delta[n] - \frac{1}{2}\delta[n-1] + \frac{1}{4}\delta[n-2]$

(d) By combining your results in part (b) with the result of Problem 10.43(b), show that for a real, even sequence, if there is a pole (zero) of $H(z)$ at $z = \rho e^{j\theta}$, then there is also a pole (zero) of $H(z)$ at $z = (1/\rho)e^{j\theta}$ and at $z = (1/\rho)e^{-j\theta}$.

10.52. Consider a sequence $x_1[n]$ with z-transform $X_1(z)$ and a sequence $x_2[n]$ with z-transform $X_2(z)$, where

$$x_2[n] = x_1[-n].$$

Show that $X_2(z) = X_1(1/z)$, and from this, show that if $X_1(z)$ has a pole (or zero) at $z = z_0$, then $X_2(z)$ has a pole (or zero) at $z = 1/z_0$.

10.53. **(a)** Carry out the proof for each of the following properties in Table 10.1:

(1) Property set forth in Section 10.5.2

(2) Property set forth in Section 10.5.3

(3) Property set forth in Section 10.5.4

(b) With $X(z)$ denoting the z-transform of $x[n]$ and R_x the ROC of $X(z)$, determine, in terms of $X(z)$ and R_x, the z-transform and associated ROC for each of the following sequences:

(1) $x^*[n]$

(2) $z_0^n x[n]$, where z_0 is a complex number

10.54. In Section 10.5.9, we stated and proved the initial-value theorem for causal sequences.

(a) State and prove the corresponding theorem if $x[n]$ is anticausal (i.e., if $x[n] = 0, n > 0$).

(b) Show that if $x[n] = 0, n < 0$, then

$$x[1] = \lim_{z \to \infty} z(X(z) - x[0]).$$

10.55. Let $x[n]$ denote a causal sequence (i.e., if $x[n] = 0, n < 0$) for which $x[0]$ is nonzero and finite.

(a) Using the initial-value theorem, show that there are no poles or zeros of $X(z)$ at $z = \infty$.

(b) Show that, as a consequence of your result in part (a), the number of poles of $X(z)$ in the finite z-plane equals the number of zeros of $X(z)$ in the finite z-plane. (The finite z-plane excludes $z = \infty$.)

10.56. In Section 10.5.7, we stated the convolution property for the z-transform. To show that this property holds, we begin with the convolution sum expressed as

$$x_3[n] = x_1[n] * x_2[n] = \sum_{k=-\infty}^{\infty} x_1[k]x_2[n-k]. \qquad \text{(P10.56–1)}$$

(a) By taking the z-transform of eq. (P10.56–1) and using eq. (10.3), show that

$$X_3(z) = \sum_{k=-\infty}^{\infty} x_1[k]\hat{X}_2(z),$$

where $\hat{X}_2(z)$ is the transform of $x_2[n-k]$.

(b) Using your result in part (a) and property 10.5.2 in Table 10.1, show that

$$X_3(z) = X_2(z) \sum_{k=-\infty}^{\infty} x_1[k]z^{-k}.$$

(c) From part (b), show that

$$X_3(z) = X_1(z)X_2(z)$$

as stated in eq. (10.81).

10.57. Let

$$X_1(z) = x_1[0] + x_1[1]z^{-1} + \cdots + x_1[N_1]z^{-N_1},$$
$$X_2(z) = x_2[0] + x_2[1]z^{-1} + \cdots + x_2[N_2]z^{-N_2}.$$

Define

$$Y(z) = X_1(z)X_2(z),$$

and let

$$Y(z) = \sum_{k=0}^{M} y[k]z^{-k}.$$

(a) Express M in terms of N_1 and N_2.
(b) Use polynomial multiplication to determine $y[0]$, $y[1]$, and $y[2]$.
(c) Use polynomial multiplication to show that, for $0 \leq k \leq M$,

$$y[k] = \sum_{m=-\infty}^{\infty} x_1[m]x_2[k-m].$$

10.58. A minimum-phase system is a system that is causal and stable and for which the inverse system is also causal and stable. Determine the necessary constraints on the location in the z-plane of the poles and zeros of the system function of a minimum-phase system.

10.59. Consider the digital filter structure shown in Figure P10.59.

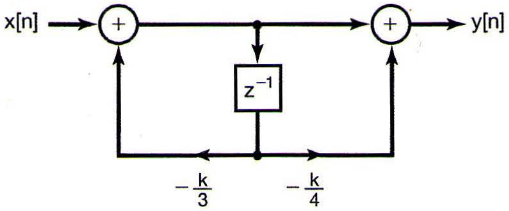


Figure P10.59

(a) Find $H(z)$ for this causal filter. Plot the pole-zero pattern and indicate the region of convergence.
(b) For what values of the k is the system stable?
(c) Determine $y[n]$ if $k = 1$ and $x[n] = (2/3)^n$ for all n.

10.60. Consider a signal $x[n]$ whose unilateral z-transform is $\mathcal{X}(z)$. Show that the unilateral z-transform of $y[n] = x[n+1]$ may be specified as

$$\mathcal{Y}(z) = z\mathcal{X}(z) - zx[0].$$

10.61. If $\mathcal{X}(z)$ denotes the unilateral z-transform of $x[n]$, determine, in terms of $\mathcal{X}(z)$, the unilateral z-transform of:
(a) $x[n+3]$ **(b)** $x[n-3]$ **(c)** $\sum_{k=-\infty}^{n} x[k]$

EXTENSION PROBLEMS

10.62. The autocorrelation sequence of a sequence $x[n]$ is defined as

$$\phi_{xx}[n] = \sum_{k=-\infty}^{\infty} x[k]x[n+k].$$

Determine the z-transform of $\phi_{xx}[n]$ in terms of the z-transform of $x[n]$.

10.63. By using the power-series expansion

$$\log(1 - w) = -\sum_{i=1}^{\infty} \frac{w^i}{i}, \quad |w| < 1,$$

determine the inverse of each of the following two z-transforms:

(a) $X(z) = \log(1 - 2z), \ |z| < \frac{1}{2}$

(b) $X(z) = \log(1 - \frac{1}{2}z^{-1}), \ |z| > \frac{1}{2}$

10.64. By first differentiating $X(z)$ and using the appropriate properties of the z-transform, determine the sequence for which the z-transform is each of the following:

(a) $X(z) = \log(1 - 2z), \ |z| < \frac{1}{2}$

(b) $X(z) = \log(1 - \frac{1}{2}z^{-1}), \ |z| > \frac{1}{2}$

Compare your results for (a) and (b) with the results obtained in Problem 10.63, in which the power-series expansion was used.

10.65. The *bilinear transformation* is a mapping for obtaining a rational z-transform $H_d(z)$ from a rational Laplace transform $H_c(s)$. This mapping has two important properties:

1. If $H_c(s)$ is the Laplace transform of a causal and stable LTI system, then $H_d(z)$ is the z-transform of a causal and stable LTI system.
2. Certain important characteristics of $|H_c(j\omega)|$ are preserved in $|H_d(e^{j\omega})|$.

In this problem, we illustrate the second of these properties for the case of all-pass filters.

(a) Let

$$H_c(s) = \frac{a - s}{s + a},$$

where a is real and positive. Show that

$$|H_c(j\omega)| = 1.$$

(b) Let us now apply the bilinear transformation to $H_c(s)$ in order to obtain $H_d(z)$. That is,

$$H_d(z) = H_c(s)\Big|_{s = \frac{1-z^{-1}}{1+z^{-1}}}.$$

Show that $H_d(z)$ has one pole (which is inside the unit circle) and one zero (which is outside the unit circle).

(c) For the system function $H_d(z)$ derived in part (b), show that $|H_d(e^{j\omega})| = 1$.

10.66. The bilinear transformation, introduced in the previous problem, may also be used to obtain a discrete-time filter, the magnitude of whose frequency response is similar to the magnitude of the frequency response of a given continuous-time lowpass filter. In this problem, we illustrate the similarity through the example of a continuous-time second-order Butterworth filter with system function $H_c(s)$.

(a) Let

$$H_d(z) = H_c(s)\big|_{s=\frac{1-z^{-1}}{1+z^{-1}}}.$$

Show that

$$H_d(e^{j\omega}) = H_c\left(j\tan\frac{\omega}{2}\right).$$

(b) Given that

$$H_c(s) = \frac{1}{(s+e^{j\pi/4})(s+e^{-j\pi/4})}$$

and that the corresponding filter is causal, verify that $H_c(0) = 1$, that $|H_c(j\omega)|$ decreases monotonically with increasing positive values of ω, that $|H_c(j)|^2 = 1/2$ (i.e., that $\omega_c = 1$ is the half-power frequency), and that $H_c(\infty) = 0$.

(c) Show that if the bilinear transformation is applied to $H_c(s)$ of part (b) in order to obtain $H_d(z)$, then the following may be asserted about $H_d(z)$ and $H_d(e^{j\omega})$:

1. $H_d(z)$ has only two poles, both of which are *inside* the unit circle.

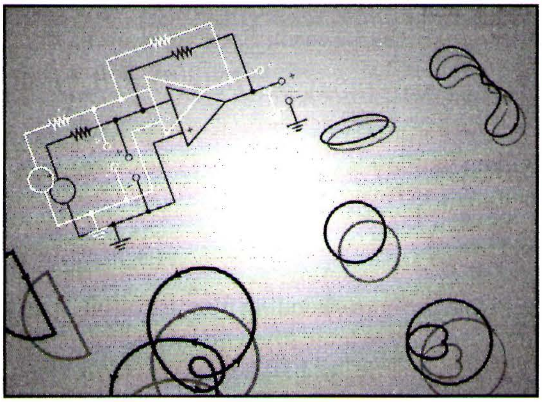
2. $H_d(e^{j0}) = 1$.

3. $|H_d(e^{j\omega})|$ decreases monotonically as ω goes from 0 to π.

4. The half-power frequency of $H_d(e^{j\omega})$ is $\pi/2$.

11

LINEAR FEEDBACK SYSTEMS



11.0 INTRODUCTION

It has long been recognized that in many situations there are particular advantages to be gained by using feedback—that is, by using the output of a system to control or modify the input. For example, it is common in electromechanical systems, such as a motor whose shaft position is to be maintained at a constant angle, to measure the error between the desired and the true position and to use this error in the form of a signal to turn the shaft in the appropriate direction. This is illustrated in Figure 11.1, where we have depicted the use of a dc motor for the accurate pointing of a telescope. In Figure 11.1(a) we have indicated pictorially what such a system would look like, where $v(t)$ is the input voltage to the motor and $\theta(t)$ is the angular position of the telescope platform. The block diagram for the motor-driven pointing system is shown in Figure 11.1(b). A feedback system for controlling the position of the telescope is illustrated in Figure 11.1(c), and a block diagram equivalent to this system is shown in Figure 11.1(d). The external, or *reference,* input to this feedback system is the desired shaft angle θ_D. A potentiometer is used to convert the angle into a voltage $K_1\theta_D$ proportional to θ_D. Similarly, a second potentiometer produces a voltage $K_1\theta(t)$ proportional to the actual platform angle. These two voltages are compared, producing an error voltage $K_1\left(\theta_D - \theta(t)\right)$, which is amplified and then used to drive the electric motor.

Figure 11.1 suggests two different methods for pointing the telescope. One of these is the feedback system of Figures 11.1(c) and (d). Here, the input that we must provide is the desired reference angle θ_D. Alternatively, if the initial angle, the desired angle, and the detailed electrical and mechanical characteristics of the motor-shaft assembly were known exactly, we could specify the precise history of the input voltage $v(t)$ that would first accelerate and then decelerate the shaft, bringing the platform to a stop at the desired

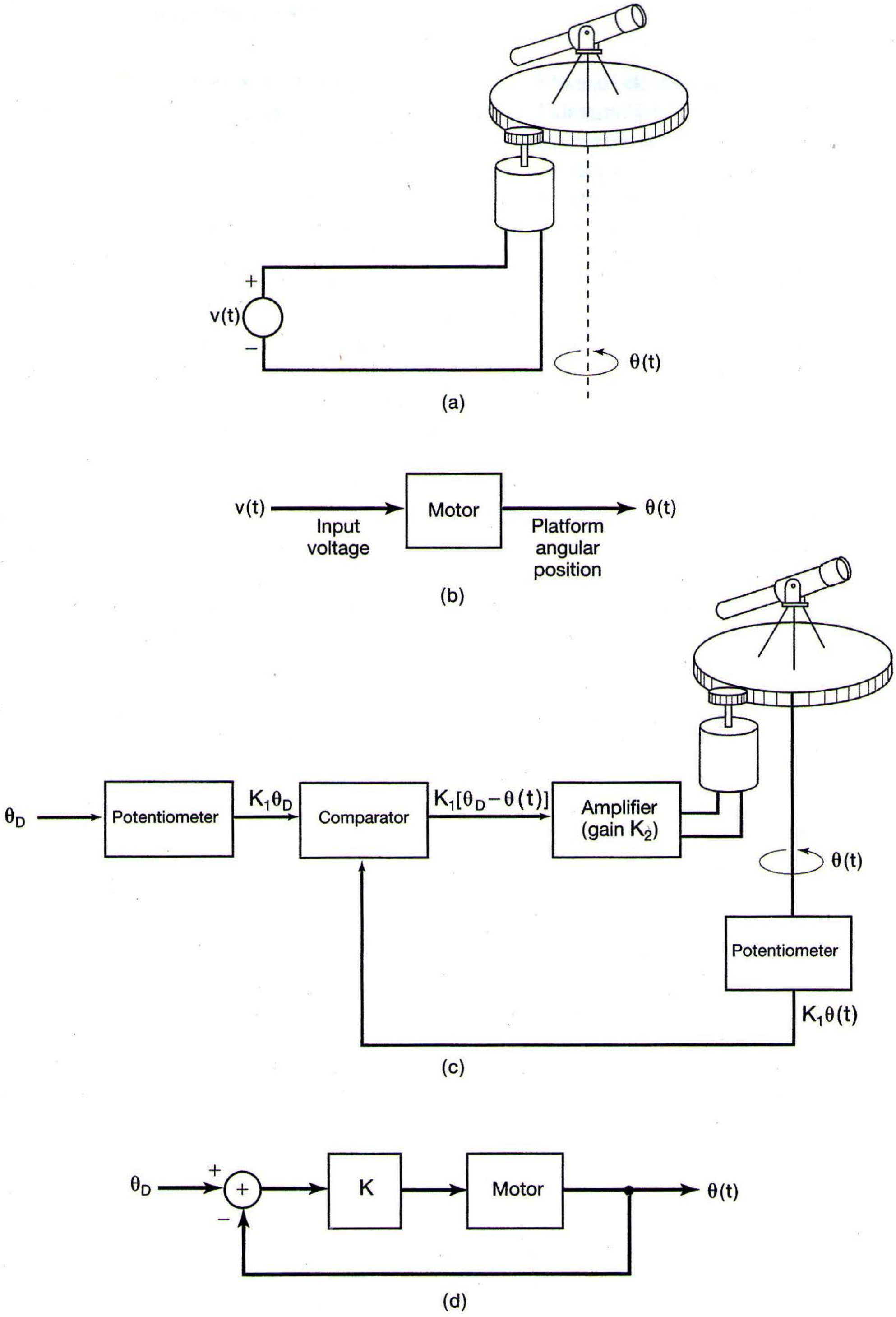

+
v(t)
−
θ(t)
(a)
v(t)
Input voltage
Motor
Platform angular position
θ(t)
(b)
θ_D
Potentiometer
K₁θ_D
Comparator
K₁[θ_D − θ(t)]
Amplifier (gain K₂)
θ(t)
Potentiometer
K₁θ(t)
(c)
θ_D
+
+
−
K
Motor
θ(t)
(d)

Figure 11.1 Use of feedback to control the angular position of a telescope: (a) dc motor-driven telescope platform; (b) block diagram of the system in (a); (c) feedback system for pointing the telescope; (d) block diagram of the system in (c) (here, $K = K_1 K_2$).

position without the use of feedback, as in Figures 11.1(a) and (b). A system operating in accordance with Figures 11.1(a) and (b) is typically referred to as an *open-loop system,* in contrast to the *closed-loop system* of Figures 11.1(c) and (d). In a practical environment, there are clear advantages to controlling the motor-shaft angle with the closed-loop system rather than with the open-loop system. For example, in the closed-loop system, when the shaft has been rotated to the correct position, any disturbance from this position will be sensed, and the resulting error will be used to provide a correction. In the open-loop system, there is no mechanism for providing a correction. As another advantage of the closed-loop system, consider the effect of errors in modeling the characteristics of the motor-shaft assembly. In the open-loop system, a precise characterization of the system is required to design the correct input. In the closed-loop system, the input is simply the desired shaft angle and does not require precise knowledge of the system. This insensitivity of the closed-loop system to disturbances and to imprecise knowledge of the system are two important advantages of feedback.

The control of an electric motor is just one of a great many examples in which feedback plays an important role. Similar uses of feedback can be found in a wide variety of applications, such as chemical process control, automotive fuel systems, household heating systems, and aerospace systems, to name just a few. In addition, feedback is also present in many biological processes and in the control of human motion. For example, when a person reaches for an object, it is usual during the reaching process to monitor visually the distance between the hand and the object so that the velocity of the hand can be smoothly decreased as the distance (i.e., the error) between the hand and the object decreases. The effectiveness of using the system output (hand position) to control the input is clearly demonstrated by alternatively reaching with and without the use of visual feedback.

In addition to its use in providing an error-correcting mechanism that can reduce sensitivity to disturbances and to errors in the modeling of the system that is to be controlled, another important characteristic of feedback is its potential for stabilizing a system that is inherently unstable. Consider the problem of trying to balance a broomstick in the palm of the hand. If the hand is held stationary, small disturbances (such as a slight breeze or inadvertent motion of the hand) will cause the broom to fall over. Of course, if one knows exactly what disturbances will occur, and if one can control the motion of the hand perfectly, it is possible to determine in advance how to move the hand to balance the broom. This is clearly unrealistic; however, by always moving the hand in the direction in which the broom is falling, the broom can be balanced. This, of course, *requires* feedback in order to sense the direction in which the broom is falling. A second example that is closely related to the balancing of a broom is the problem of controlling a so-called inverted pendulum, which is illustrated in Figure 11.2. As shown, an inverted pendulum consists of a thin rod with a weight at the top. The bottom of the rod is mounted on a cart that can move in either direction along a track. Again, if the cart is kept stationary, the inverted pendulum

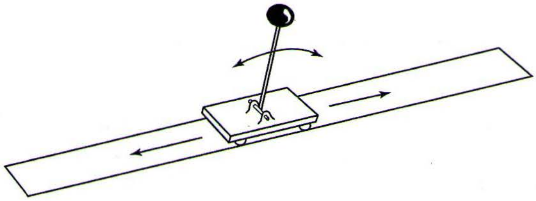


Figure 11.2 An inverted pendulum.

will topple over. The problem of stabilizing the pendulum is one of designing a feedback system that will move the cart to keep the pendulum vertical. This example is examined in Problem 11.56. A third example, which again bears some similarity to the balancing of a broom, is the problem of controlling the trajectory of a rocket. In this case, much as the movement of the hand is used to compensate for disturbances in the position of the broom, the direction of the thrust of the rocket is used to correct for changes in aerodynamic forces and wind disturbances that would otherwise cause the rocket to deviate from its course. Again, feedback is important, because these forces and disturbances are never precisely known in advance.

The preceding examples provide some indication of why feedback may be useful. In the next two sections we introduce the basic block diagrams and equations for linear feedback systems and discuss in more detail a number of applications of feedback and control, both in continuous time and in discrete time. We also point out how feedback can have harmful as well as useful effects. These examples of the uses and effects of feedback will give us some insight into how changes in the parameters in a feedback control system lead to changes in the behavior of the system. Understanding this relationship is essential in designing feedback systems that have certain desirable characteristics. With this material as background, we will then develop, in the remaining sections of the chapter, several specific techniques that are of significant value in the analysis and design of continuous-time and discrete-time feedback systems.

11.1 LINEAR FEEDBACK SYSTEMS

The general configuration of a continuous-time LTI feedback system is shown in Figure 11.3(a) and that of a discrete-time LTI feedback system in Figure 11.3(b). Because of

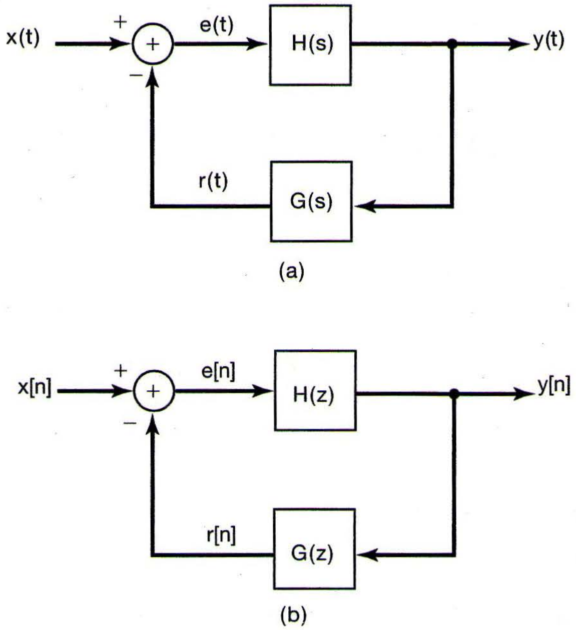


Figure 11.3 Basic feedback system configurations in (a) continuous time and (b) discrete time.

the typical applications in which feedback is utilized, it is natural to restrict the systems in these figures to be causal. This will be our assumption throughout the chapter. In that case, the system functions in Figure 11.3 can be interpreted either as unilateral or as bilateral transforms, and, as a consequence of causality, the ROC's associated with them will always be to the right of the rightmost pole for Laplace transforms and outside the outermost pole for z-transforms.

It should also be noted that the convention used in Figure 11.3(a) is that $r(t)$, the signal fed back, is subtracted from the input $x(t)$ to form $e(t)$. The identical convention is adopted in discrete time. Historically, this convention arose in tracking-system applications, where $x(t)$ represented a desired command and $e(t)$ represented the error between the command and the actual response $r(t)$. This was the case, for example, in our discussion of the pointing of a telescope. In more general feedback systems, $e(t)$ and $e[n]$, the discrete-time counterpart of $e(t)$, may not correspond to or be directly interpretable as error signals.

The system function $H(s)$ in Figure 11.3(a) or $H(z)$ in Figure 11.3(b) is referred to as the *system function of the forward path* and $G(s)$ or $G(z)$ as the *system function of the feedback path.* The system function of the overall system of Figure 11.3(a) or (b) is referred to as the *closed-loop system function* and will be denoted by $Q(s)$ or $Q(z)$. In Sections 9.8.1 and 10.8.1, we derived expressions for the system functions of feedback interconnections of LTI systems. Applying these results to the feedback systems of Figure 11.3, we obtain

$$Q(s) = \frac{Y(s)}{X(s)} = \frac{H(s)}{1+G(s)H(s)}, \tag{11.1}$$

$$Q(z) = \frac{Y(z)}{X(z)} = \frac{H(z)}{1+G(z)H(z)}. \tag{11.2}$$

Equations (11.1) and (11.2) represent the fundamental equations for the study of LTI feedback systems. In the following sections, we use these equations as the basis for gaining insight into the properties of feedback systems and for developing several tools for their analysis.

11.2 SOME APPLICATIONS AND CONSEQUENCES OF FEEDBACK

In the introduction, we provided a brief, intuitive look at some of the properties and uses of feedback systems. In this section, we examine a number of the characteristics and applications of feedback in somewhat more quantitative terms, using the basic feedback equations (11.1) and (11.2) as a starting point. Our purpose is to provide an introduction to and an appreciation for the applications of feedback, rather than to develop any of these applications in detail. In the sections that follow, we focus in more depth on several specific techniques for analyzing feedback systems that are useful in a wide range of problems, including many of the applications that we are about to describe.

11.2.1 Inverse System Design

In some applications, one would like to synthesize the inverse of a given continuous-time system. Suppose that this system has system function $P(s)$, and consider the feedback system shown in Figure 11.4. Applying equation (11.1) with $H(s) = K$ and $G(s) = P(s)$,

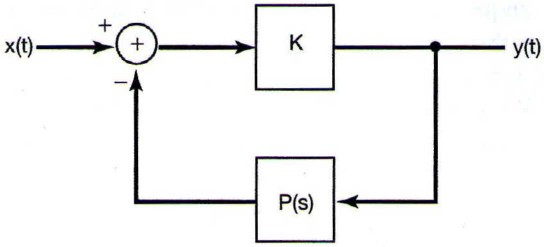


Figure 11.4 Form of a feedback system used in implementing the inverse of the system with system function $P(s)$.

we find that the closed-loop system function is

$$Q(s) = \frac{K}{1 + KP(s)}. \tag{11.3}$$

If the gain K is sufficiently large so that $KP(s) \gg 1$, then

$$Q(s) \simeq \frac{1}{P(s)}, \tag{11.4}$$

in which case the feedback system approximates the inverse of the system with system function $P(s)$.

It is important to note that the result in eq. (11.4) requires that the gain K be sufficiently high, but is otherwise not dependent on the precise value of the gain. Operational amplifiers are one class of devices that provide this kind of gain and are widely used in feedback systems. One common application of the inversion inherent in eq. (11.4) is in the implementation of integrators. A capacitor has the property that its current is proportional to the derivative of the voltage. By inserting a capacitor in the feedback path around an operational amplifier, the differentiation property of the capacitor is inverted to provide integration. This specific application is explored in more detail in Problems 11.50–11.52.

Although our discussion is for the most part restricted to linear systems, it is worth pointing out that this same basic approach is commonly used in inverting a nonlinearity. For example, systems for which the output is the logarithm of the input are commonly implemented by utilizing the exponential current-voltage characteristics of a diode as feedback around an operational amplifier. This is explored in more detail in Problem 11.53.

11.2.2 Compensation for Nonideal Elements

Another common use of feedback is to correct for some of the nonideal properties of the open-loop system. For example, feedback is often used in the design of amplifiers to provide constant-gain amplification in a given frequency band, and in fact, it is this application, pioneered by H. S. Black at Bell Telephone Laboratories in the 1920s, that is generally considered to have been the catalyst for the development of feedback control as a practical and useful system design methodology.

Specifically, consider an open-loop frequency response $H(j\omega)$ which provides amplification over the specified frequency band, but which is not constant over that range. For example, operational amplifiers or the vacuum tube amplifiers of concern to Black and his colleagues typically provide considerable, but not precisely controlled, amplification.

While such devices can provide raw amplification levels of several orders of magnitude, the price one pays for this includes uncertain levels of amplification that can fluctuate with frequency, time, temperature, etc., and that can also introduce unwanted phase and nonlinear distortions. What Black proposed was placing such a powerful, but uncertain and erratic, amplifier in a feedback loop as in Figure 11.3(a) with $G(s)$ chosen to be constant, i.e., $G(s) = K$. In this case, assuming the closed-loop system is stable, its frequency response is

$$Q(j\omega) = \frac{H(j\omega)}{1 + KH(j\omega)}. \tag{11.5}$$

If, over the specified frequency range,

$$|KH(j\omega)| \gg 1, \tag{11.6}$$

then

$$Q(j\omega) \simeq \frac{1}{K}. \tag{11.7}$$

That is, the closed-loop frequency response is constant, as desired. This of course assumes that the system in the feedback path can be designed so that its frequency response $G(j\omega)$ has a constant gain K over the desired frequency band, which is precisely what we assumed we could *not* ensure for $H(j\omega)$. The difference between the requirement on $H(j\omega)$ and that on $G(j\omega)$, however, is that $H(j\omega)$ must provide amplification, whereas, from eq. (11.7), we see that for the overall closed-loop system to provide a gain greater than unity, K must be less than 1. That is, $G(j\omega)$ must be an attenuator over the specified range of frequencies. In general, an attenuator with approximately flat frequency characteristics is considerably easier to realize than an amplifier with approximately flat frequency response (since an attenuator can be constructed from passive elements).

The use of feedback to flatten the frequency response incurs some cost, however, and it is this fact that led to the considerable skepticism with which Black's idea was met. In particular, from eqs. (11.6) and (11.7), we see that

$$|H(j\omega)| \gg \frac{1}{K} \simeq Q(j\omega), \tag{11.8}$$

so that the closed-loop gain $1/K$ will be substantially less than the open-loop gain $|H(j\omega)|$. This apparently significant loss of gain, attributable to what Black referred to as *degenerative* or negative feedback, was initially viewed as a serious weakness in his negative-feedback amplifier. Indeed, the effect had been known for many years and had led to the conviction that negative feedback was not a particularly useful mechanism. However, Black pointed out that what one gave up in overall gain was often more than offset by the reduced sensitivity of the overall closed-loop amplifier: The closed-loop system function is essentially equal to eq. (11.7), independently of variations in $H(j\omega)$, as long as $|H(j\omega)|$ is large enough. Thus, if the open-loop amplifier is initially designed with considerably more gain than is actually needed, the closed-loop amplifier will provide the desired levels of amplification with greatly reduced sensitivity. This concept and its application to extending the bandwidth of an amplifier are explored in Problem 11.49.

11.2.3 Stabilization of Unstable Systems

As mentioned in the introduction, one use of feedback systems is to stabilize systems that, without feedback, are unstable. Examples of this kind of application include the control of the trajectory of a rocket, the regulation of nuclear reactions in a nuclear power plant, the stabilization of an aircraft, and the natural and regulatory control of animal populations.

To illustrate how feedback can be used to stabilize an unstable system, let us consider a simple first-order continuous-time system with

$$H(s) = \frac{b}{s-a}. \tag{11.9}$$

With $a > 0$, the system is unstable. Choosing the system function $G(s)$ to be a constant gain K, we see that the closed-loop system function in eq. (11.1) becomes

$$\begin{aligned} Q(s) &= \frac{H(s)}{1+KH(s)} \\ &= \frac{b}{s-a+Kb}. \end{aligned} \tag{11.10}$$

The closed-loop system will be stable if the pole is moved into the left half of the s-plane. This will be the case if

$$Kb > a. \tag{11.11}$$

Thus, we can stabilize the system with a constant gain in the feedback loop if that gain is chosen to satisfy eq. (11.11). This type of feedback system is referred to as a *proportional feedback system,* since the signal that is fed back is proportional to the output of the system.

As another example, consider the second-order system

$$H(s) = \frac{b}{s^2+a}. \tag{11.12}$$

If $a > 0$, the system is an oscillator (i.e., $H(s)$ has its poles on the $j\omega$-axis), and the impulse response of the system is sinusoidal. If $a < 0$, $H(s)$ has one pole in the left-half plane and one in the right-half plane. Thus, in either case, the system is unstable. In fact, as considered in Problem 11.56, the system function given in eq. (11.12) with $a < 0$ can be used to model the dynamics of the inverted pendulum described in the introduction.

Let us first consider the use of proportional feedback for this second-order system; that is, we take

$$G(s) = K. \tag{11.13}$$

Substituting into eq. (11.1), we obtain

$$Q(s) = \frac{b}{s^2+(a+Kb)}. \tag{11.14}$$

In our discussion of second-order systems in Chapter 6, we considered a transfer function of the form

$$\frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}. \tag{11.15}$$

For such a system to be stable, ω_n must be real and positive (i.e., $\omega_n^2 > 0$), and ζ must be positive (corresponding to positive damping). From eqs. (11.14) and (11.15), it follows that with proportional feedback we can only influence the value of ω_n^2, and consequently, we cannot stabilize the system because we cannot introduce any damping. To suggest a type of feedback that can be used to stabilize this system, recall the mass-spring-dashpot mechanical system described in our examination of second-order systems in Section 6.5.2. We saw that damping in that system was the result of the inclusion of a dashpot, which provided a restoring force proportional to the *velocity* of the mass. This suggests that we consider *proportional-plus-derivative* feedback, that is, a $G(s)$ of the form

$$G(s) = K_1 + K_2 s, \tag{11.16}$$

which yields

$$Q(s) = \frac{b}{s^2 + bK_2 s + (a + K_1 b)}. \tag{11.17}$$

The closed-loop poles will be in the left-half plane, and hence, the closed-loop system will be stable as long as we choose K_1 and K_2 to guarantee that

$$bK_2 > 0, \qquad a + K_1 b > 0. \tag{11.18}$$

The preceding discussion illustrates how feedback can be used to stabilize continuous-time systems. The stabilization of unstable systems is an important application of feedback for discrete-time systems as well. Examples of discrete-time systems that are unstable in the absence of feedback are models of population growth. To illustrate how feedback can prevent the unimpeded growth of populations, let us consider a simple model for the evolution of the population of a single species of animal. Let $y[n]$ denote the number of animals in the nth generation, and assume that without the presence of any impeding influences, the birthrate is such that the population would double each generation. In this case, the basic equation for the population dynamics of the species is

$$y[n] = 2y[n-1] + e[n], \tag{11.19}$$

where $e[n]$ represents any additions to or deletions from the population that are caused by external influences.

This population model is obviously unstable, with an impulse response that grows exponentially. However, in any ecological system, there are a number of factors that will inhibit the growth of a population. For example, limits on the food supply for the species will manifest themselves through a reduction in population growth when the number of animals becomes large. Similarly, if the species has natural enemies, it is often reasonable to assume that the population of the predators will grow when the population of the prey increases and, consequently, that the presence of natural enemies will retard population growth. In addition to natural influences such as these, there may be effects introduced by humans that are aimed at population control. For example, the food supply or the predator population may fall under human regulation. In addition, stocking lakes with fish or importing animals from other areas can be used to promote growth, and the control of hunting or fishing can also provide a regulative effect. Because all of these influences depend on the size of the population (either naturally or by design), they represent feedback effects.

Based on the preceding discussion, we can separate $e[n]$ into two parts by means of the equation

$$e[n] = x[n] - r[n], \tag{11.20}$$

where $r[n]$ represents the effect of the regulative influences described in the previous paragraph and $x[n]$ incorporates any other external effects, such as the migration of animals or natural disasters or disease. Note that we have included a minus sign in eq. (11.20). This is consistent with our convention of using negative feedback, and here it also has the physical interpretation that, since the uninhibited growth of the population is unstable, the feedback term plays the role of a *retarding* influence. To see how the population can be controlled by the presence of this feedback term, suppose that the regulative influences account for the depletion of a fixed proportion β of the population in each generation. Since, according to our model, the surviving fraction of each generation will double in size, it follows that

$$y[n] = 2(1-\beta)y[n-1] + x[n]. \tag{11.21}$$

Comparing eq. (11.21) with eqs. (11.19) and (11.20), we see that

$$r[n] = 2\beta y[n-1]. \tag{11.22}$$

The factor of 2 here represents the fact that the depletion of the present population decreases the number of births in the next generation.

This example of the use of feedback is illustrated in Figure 11.5. Here, the system function of the forward path is obtained from eq. (11.19) as

$$H(z) = \frac{1}{1-2z^{-1}}, \tag{11.23}$$

while from eq. (11.22) the system function of the feedback path is

$$G(z) = 2\beta z^{-1}. \tag{11.24}$$

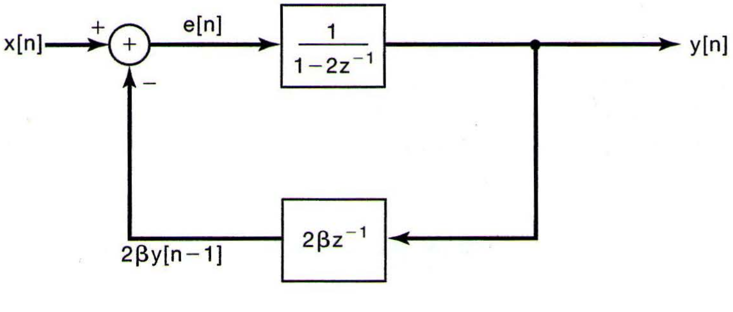


Figure 11.5 Block diagram of a simple feedback model of population dynamics.

Consequently, the closed-loop system function is

$$Q(z) = \frac{H(z)}{1 + G(z)H(z)} = \frac{1}{1 - 2(1-\beta)z^{-1}}. \tag{11.25}$$

If $\beta < 1/2$, the closed-loop system is still unstable, whereas it is stable[1] if $1/2 < \beta < 3/2$.

Clearly, this example of population growth and control is extremely simplified. For instance, the feedback model of eq. (11.22) does not account for the fact that the part of $r[n]$ which is due to the presence of natural enemies depends upon the population of the predators, which in turn has its own growth dynamics. Such effects can be incorporated by making the feedback model more complex to reflect the presence of other dynamics in an ecological system, and the resulting models for the evolution of interacting species are extremely important in ecological studies. However, even without the incorporation of these effects, the simple model that we have described here does illustrate the basic ideas of how feedback can prevent both the unlimited proliferation of a species and its extinction. In particular, we can see at an elementary level how human-induced factors can be used. For example, if a natural disaster or an increase in the population of natural enemies causes a drastic decrease in the population of a species, a tightening of limits on hunting or fishing and accelerated efforts to increase the population can be used to decrease β in order to *destabilize* the system to allow for rapid growth, until a normal-size population is again attained.

Note also that for this type of problem, it is not usually the case that one wants strict stability. If the regulating influences are such that $\beta = 1/2$, and if all other external influences are zero (i.e., if $x[n] = 0$), then $y[n] = y[n-1]$. Therefore, as long as $x[n]$ is small and averages to zero over several generations, a value of $\beta = 1/2$ will result in an essentially constant population. However, for this value of β the system is unstable, since eq. (11.21) then reduces to

$$y[n] = y[n-1] + x[n]. \tag{11.26}$$

That is, the system is equivalent to an accumulator. Thus, if $x[n]$ is a unit step, the output grows without bound. Consequently, if a steady trend is expected in $x[n]$, caused, for example, by a migration of animals into a region, a value of $\beta > 1/2$ would need to be used to stabilize the system and thus to keep the population within bounds and maintain an ecological balance.

11.2.4 Sampled-Data Feedback Systems

In addition to dealing with problems such as the one just described, discrete-time feedback techniques are of great importance in a wide variety of applications involving continuous-time systems. The flexibility of digital systems has made the implementation of *sampled-data feedback systems* an extremely attractive option. In such a system, the output of a continuous-time system is sampled, some processing is done on the resulting sequence of samples, and a discrete sequence of feedback commands is generated. This sequence is then converted to a continuous-time signal that is fed back to and subtracted from the external input to produce the actual input to the continuous-time system.

---

[1] Although, in the context of our population example, β could never exceed unity, since $\beta > 1$ corresponds to removing more than 100% of the population.

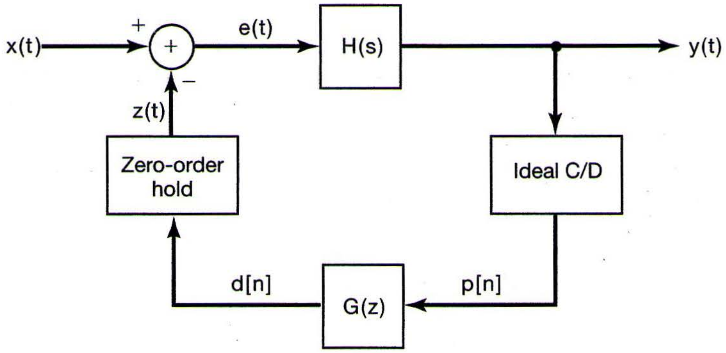
Clearly, the constraint of causality on feedback systems imposes a restriction on the process of converting the discrete-time feedback signal to a continuous-time signal (e.g., ideal lowpass filtering or any noncausal approximation of it is not allowed). One of the most widely used conversion systems is the zero-order hold (introduced in Section 7.1.2). The structure of a sampled-data feedback system involving a zero-order hold is depicted in Figure 11.6(a). In the figure, we have a continuous-time LTI system with system function $H(s)$ that is sampled to produce a discrete-time sequence

$$p[n] = y(nT). \tag{11.27}$$

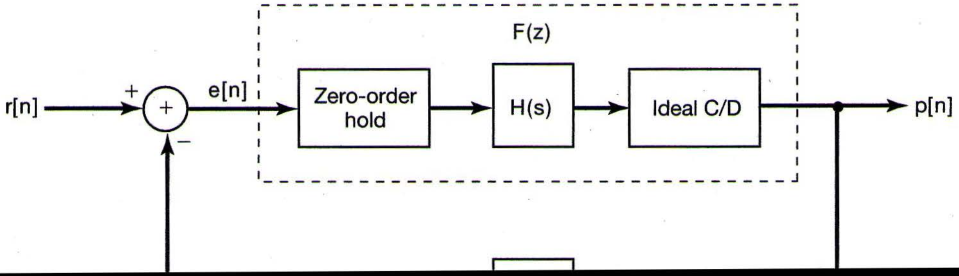
The sequence $p[n]$ is then processed by a discrete-time LTI system with system function $G(z)$, and the resulting output is put through a zero-order hold to produce the continuous-time signal

$$z(t) = d[n] \qquad \text{for} \quad nT \leq t < (n+1)T. \tag{11.28}$$

This signal is subtracted from the external input $x(t)$ to produce $e(t)$.



(a)



[redacted]

Suppose also that $x(t)$ is constant over intervals of length T. That is,

$$x(t) = r[n] \qquad \text{for} \quad nT \le t < (n+1)T, \tag{11.29}$$

where $r[n]$ is a discrete-time sequence. This is an approximation that is usually valid in practice, as the sampling rate is typically fast enough so that $x(t)$ does not change appreciably over intervals of length T. Furthermore, in many applications, the external input is itself actually generated by applying a zero-order hold operation to a discrete sequence. For example, in systems such as advanced aircraft, the external inputs represent human operator commands that are themselves first processed digitally and then converted back to continuous-time input signals. Because the zero-order hold is a linear operation, the feedback system of Figure 11.6(a) when $x(t)$ is given by eq. (11.29) is equivalent to the system of Figure 11.6(b).

As shown in Problem 11.60, the discrete-time system with input $e[n]$ and output $p[n]$ is an LTI system with system function $F(z)$ that is related to the continuous-time system function $H(s)$ by means of a *step-invariant* transformation. That is, if $s(t)$ is the step response of the continuous-time system, then the step response $q[n]$ of the discrete-time system consists of equally spaced samples of $s(t)$. Mathematically,

$$q[n] = s(nT) \qquad \text{for all } n. \tag{11.30}$$

Once we have determined $F(z)$, we have a completely discrete-time feedback system model (Figure 11.6(b)) exactly capturing the behavior of the continuous-time feedback system (Figure 11.6(a)) at the sampling instants $t = nT$, and we can then consider designing the feedback system function $G(z)$ to achieve our desired objectives. An example of designing such a sampled-data feedback system to stabilize an unstable continuous-time system is examined in detail in Problem 11.60.

11.2.5 Tracking Systems

As mentioned in Section 11.0, one of the important applications of feedback is in the design of systems in which the objective is to have the output track or follow the input. There is a broad range of problems in which tracking is an important component. For example, the telescope-pointing problem discussed in Section 11.0 is a tracking problem: The feedback system of Figures 11.1(c) and (d) has as its input the desired pointing angle, and the purpose of the feedback loop is to provide a mechanism for driving the telescope to follow the input. In airplane autopilots the input is the desired flight path of the vehicle, and the autopilot feedback system uses the aircraft control surfaces (rudder, ailerons, and elevator) and thrust control in order to keep the aircraft on the prescribed course.

To illustrate some of the issues that arise in the design of tracking systems, consider the discrete-time feedback system depicted in Figure 11.7(a). The examination of discrete-time tracking systems of this form often arises in analyzing the characteristics of sampled-data tracking systems for continuous-time applications. One example of such a system is a digital autopilot. In Figure 11.7(a), $H_p(z)$ denotes the system function of the system whose output is to be controlled. This system is often referred to as the *plant,* a term that can be traced to applications such as the control of power plants, heating systems, and chemical-processing plants. The system function $H_c(z)$ represents a compensator, which is the element to be designed. Here, the input to the compensator is the tracking error—

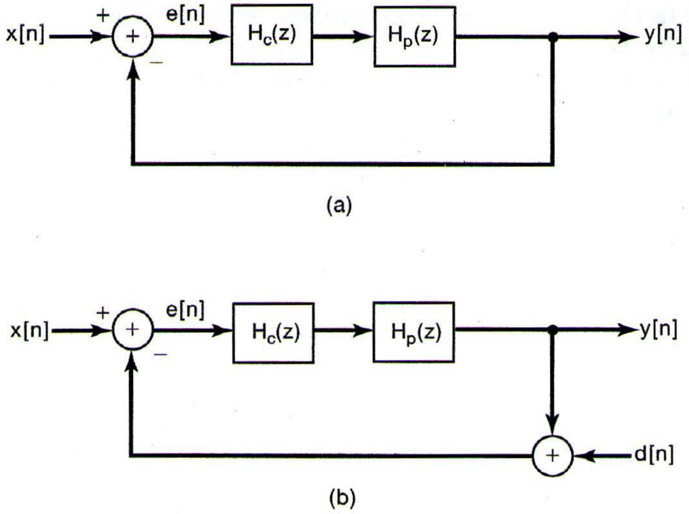


Figure 11.7 (a) Discrete-time tracking system; (b) tracking system of (a) with a disturbance $d[n]$ in the feedback path accounting for the presence of measurement errors.

that is, the difference $e[n]$ between the input $x[n]$ and the output $y[n]$. The output of the compensator is the input to the plant (for example, the actual voltage applied to the motor in the feedback system of Figures 11.1(c) and (d) or the actual physical input to the drive system of the rudder of an aircraft).

To simplify notation, let $H(z) = H_c(z)H_p(z)$. In this case, the application of eq. (11.2) yields the relationship

$$Y(z) = \frac{H(z)}{1 + H(z)}X(z). \tag{11.31}$$

Also, since $Y(z) = H(z)E(z)$, it follows that

$$E(z) = \frac{1}{1 + H(z)}X(z), \tag{11.32}$$

or, specializing to $z = e^{j\omega}$, we obtain

$$E(e^{j\omega}) = \frac{1}{1 + H(e^{j\omega})}X(e^{j\omega}). \tag{11.33}$$

Equation (11.33) provides us with some insight into the design of tracking systems. Specifically, for good tracking performance, we would like $e[n]$ or, equivalently, $E(e^{j\omega})$ to be small. That is,

$$\frac{1}{1 + H(e^{j\omega})}X(e^{j\omega}) \simeq 0. \tag{11.34}$$

Consequently, for that range of frequencies for which $X(e^{j\omega})$ is nonzero, we would like $|H(e^{j\omega})|$ to be large. Thus, we have one of the fundamental principles of feedback system design: Good tracking performance requires a large gain. This desire for a large gain, however, must typically be tempered, for several reasons. One reason is that if the gain is *too* large, the closed-loop system may have undesirable characteristics (such as too little damping) or might in fact become unstable. This possibility is discussed in the next section and is also addressed by the methods developed in subsequent sections.

In addition to the issue of stability, there are other reasons for wanting to limit the gain in a tracking system. For example, in implementing such a system, we must measure the output $y[n]$ in order to compare it to the command input $x[n]$, and any measuring device used will have inaccuracies and error sources (such as thermal noise in the electronics of the device). In Figure 11.7(b), we have included these error sources in the form of a disturbance input $d[n]$ in the feedback loop. Some simple system function algebra yields the following relationship between $Y(z)$ and the transforms $X(z)$ and $D(z)$ of $x[n]$ and $d[n]$:

$$Y(z) = \left[\frac{H(z)}{1+H(z)}X(z)\right] - \left[\frac{H(z)}{1+H(z)}D(z)\right]. \tag{11.35}$$

From this expression, we see that in order to minimize the influence of $d[n]$ on $y[n]$, we would like $H(z)$ to be small so that the second term on the right-hand side of eq. (11.35) is small.

From the preceding development, we see that the goals of tracking and of minimizing the effect of measurement errors are conflicting, and one must take this into account in coming up with an acceptable system design. In general, the design depends on more detailed information concerning the characteristics of the input $x[n]$ and the disturbance $d[n]$. For example, in many applications $x[n]$ has a significant amount of its energy concentrated at low frequencies, while measurement error sources such as thermal noise have a great deal of energy at high frequencies. Consequently, one usually designs the compensator $H_c(z)$ so that $|H(e^{j\omega})|$ is large at low frequencies and is small for ω near $\pm\pi$.

There are a variety of other issues that one must consider in designing tracking systems, such as the presence of disturbances at other points in the feedback loop. (For example, the effect of wind on the motion of an aircraft must be taken into account in designing an autopilot.) The methods of feedback system analysis introduced in this chapter provide the necessary tools for examining each of these issues. In Problem 11.57, we use some of these tools to investigate several other aspects of the problem of designing tracking systems.

11.2.6 Destabilization Caused by Feedback

As well as having many applications, feedback can have undesirable effects and can in fact cause instability. For example, consider the telescope-pointing system illustrated in Figure 11.1. From the discussion in the preceding section, we know that it would be desirable to have a large amplifier gain in order to achieve good performance in tracking the desired pointing angle. On the other hand, as we increase the gain, we are likely to obtain faster tracking response at the expense of a reduction in system damping, resulting in significant overshoot and ringing in response to changes in the desired angle. Furthermore, instability can result if the gain is increased too much.

Another common example of the possible destabilizing effect of feedback is feedback in audio systems. Consider the situation depicted in Figure 11.8(a). Here, a loudspeaker produces an audio signal that is an amplified version of the sounds picked up by a microphone. Note that in addition to other audio inputs, the sound coming from the speaker itself may be sensed by the microphone. How strong this particular signal is depends upon

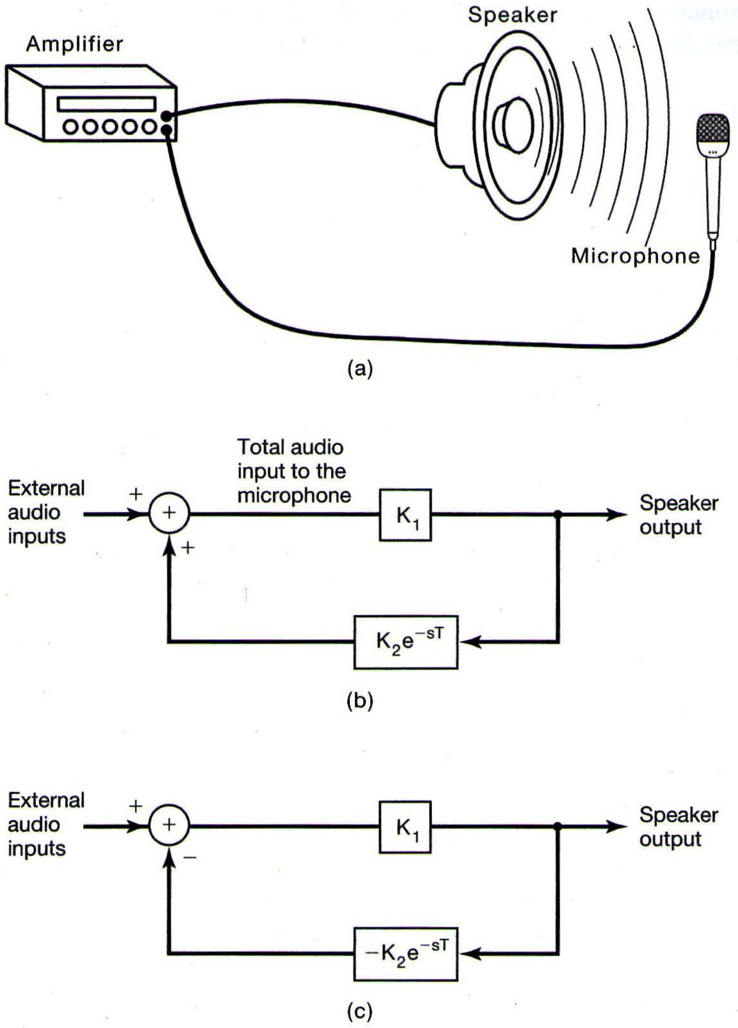

Speaker
Amplifier
Microphone
(a)
Total audio input to the microphone
External audio inputs
+
+
+
K1
K2e−sT
Speaker output
(b)
External audio inputs
+
+
−
K1
−K2e−sT
Speaker output
(c)

Figure 11.8 (a) Pictorial representation of the phenomenon of audio feedback; (b) block diagram representation of (a); (c) block diagram in (b) redrawn as a negative feedback system. (*Note:* e^{-sT} is the system function of a T-second time delay.)

the distance between the speaker and the microphone. Specifically, because of the attenuating properties of air, the strength of the signal reaching the microphone from the speaker decreases as the distance between the speaker and the microphone increases. In addition, due to the finite speed of propagation of sound waves, there is time delay between the signal produced by the speaker and that sensed by the microphone.

This audio feedback system is represented in block diagram form in Figure 11.8(b). Here, the constant K_2 in the feedback path represents the attenuation, and T is the propagation delay. The constant K_1 is the amplifier gain. Also, note that the output from the feedback path is *added* to the external input. This is an example of *positive feedback.* As discussed at the beginning of the section, the use of a negative sign in the definition of the basic feedback system of Figure 11.3 is purely conventional, and positive and negative feedback systems can be analyzed using the same tools. For example, as illustrated in Figure 11.8(c), the feedback system of Figure 11.8(b) can be written as a negative feedback system by adding a minus sign to the feedback-path system function. From this figure and from eq. (11.1), we can determine the closed-loop system function:

$$Q(s) = \frac{K_1}{1 - K_1 K_2 e^{-sT}}. \tag{11.36}$$

Later we will return to this example, and, using a technique that we will develop in Section 11.3, we will show that the system of Figure 11.8 is unstable if

$$K_1 K_2 > 1. \tag{11.37}$$

Since the attenuation due to the propagation of sound through the air decreases (i.e., K_2 *increases*) as the distance between the speaker and the microphone decreases, if the microphone is placed too close to the speaker, so that eq. (11.37) is satisfied, the system will be unstable. The result of this instability is an excessive amplification and distortion of audio signals.

It is interesting to note that positive, or what Black referred to as *regenerative,* feedback had also been known for some time before he invented his negative feedback amplifier and, ironically, had been viewed as a very useful mechanism (in contrast to the skeptical view of negative feedback). Indeed, positive feedback can be useful. For example, it was already known in the 1920s that the destabilizing influence of positive feedback could be used to generate oscillating signals. This use of positive feedback is illustrated in Problem 11.54.

In this section, we have described a number of the applications of feedback. These and others, such as the use of feedback in the implementation of recursive discrete-time filters (see Problem 11.55), are considered in more detail in the problems at the end of the chapter. From our examination of the uses of feedback and the possible stabilizing and destabilizing effects that it can have, it is clear that some care must be taken in designing and analyzing feedback systems to ensure that the closed-loop system behaves in a desirable fashion. Specifically, in Sections 11.2.3 and 11.2.6, we have seen several examples of feedback systems in which the characteristics of the closed-loop system can be significantly altered by changing the values of one or two parameters in the feedback system. In the remaining sections of this chapter, we develop several techniques for analyzing the effect of changes in such parameters on the closed-loop system and for designing systems to meet desired objectives such as stability, adequate damping, etc.

11.3 ROOT-LOCUS ANALYSIS OF LINEAR FEEDBACK SYSTEMS

As we have seen in a number of the examples and applications we have discussed, a useful type of feedback system is that in which the system has an adjustable gain K associated with it. As this gain is varied, it is of interest to examine how the poles of the closed-loop system change, since the locations of these poles tell us a great deal about the behavior of the system. For example, in stabilizing an unstable system, the adjustable gain is used to move the poles into the left-half plane for a continuous-time system or inside the unit circle for a discrete-time system. In addition, in Problem 11.49, we show that feedback can be used to broaden the bandwidth of a first-order system by moving the pole so as to decrease the time constant of the system. Furthermore, just as feedback can be used to relocate the poles to improve system performance, as we saw in Section 11.2.6, there is the potential danger that with an improper choice of feedback a stable system can be destabilized, which is usually undesirable.

In this section, we discuss a particular method for examining the locus (i.e., the path) in the complex plane of the poles of the closed-loop system as an adjustable gain is varied. The procedure, referred to as the *root-locus method,* is a graphical technique for plotting the closed-loop poles of a rational system function $Q(s)$ or $Q(z)$ as a function of the value of the gain. The technique works in an identical manner for both continuous-time and discrete-time systems.

11.3.1 An Introductory Example

To illustrate the basic nature of the root-locus method for analyzing a feedback system, let us reexamine the discrete-time example considered in the preceding section and specified by the system functions

$$\text{[eq. (11.23)]} \qquad H(z) = \frac{1}{1 - 2z^{-1}} = \frac{z}{z-2} \tag{11.38}$$

and

$$\text{[eq. (11.24)]} \qquad G(z) = 2\beta z^{-1} = \frac{2\beta}{z}, \tag{11.39}$$

where β now is viewed as an adjustable gain. Then, as we noted earlier, the closed-loop system function is

$$\text{[eq. (11.25)]} \qquad Q(z) = \frac{1}{1 - 2(1-\beta)z^{-1}} = \frac{z}{z - 2(1-\beta)}. \tag{11.40}$$

In this example, it is straightforward to identify the closed-loop pole as being located at $z = 2(1-\beta)$. In Figure 11.9(a), we have plotted the locus of the pole for the system as β varies from 0 to $+\infty$. In part (b) of the figure, we have plotted the locus as β varies from 0 to $-\infty$. In each plot, we have indicated the point $z = 2$, which is the open-loop pole [i.e., it is the pole of $Q(z)$ for $\beta = 0$]. As β increases from 0, the pole moves to the left of the point $z = 2$ along the real axis, and we have indicated this by including an arrow on the thick line to show how the pole changes as β is increased. Similarly, for $\beta < 0$, the pole of $Q(z)$ moves to the *right* of $z = 2$, and the direction of the arrow in Figure 11.9(b) indicates how the pole changes as the magnitude of β increases. For $1/2 < \beta < 3/2$, the pole lies inside the unit circle, and thus, the system is stable.

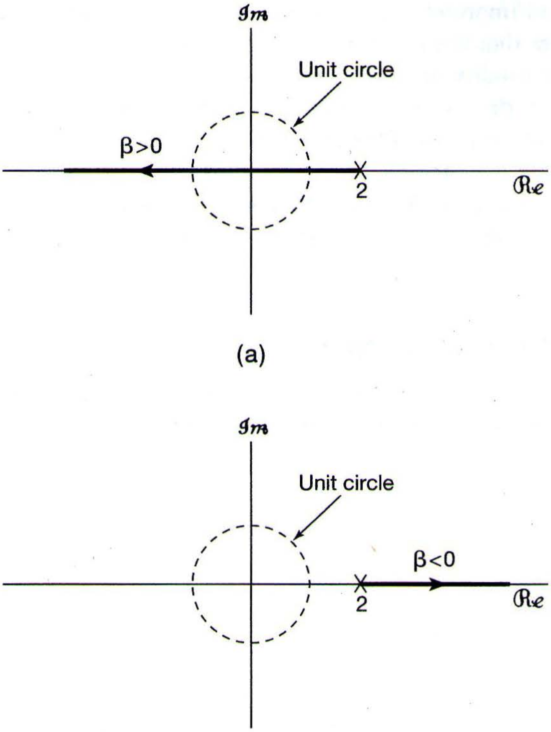
As a second example, consider a continuous-time feedback system with

$$H(s) = \frac{s}{s-2} \tag{11.41}$$

and

$$G(s) = \frac{2\beta}{s}, \tag{11.42}$$

where β again represents the adjustable gain. Since $H(s)$ and $G(s)$ in this example are algebraically identical to $H(z)$ and $G(z)$, respectively, in the preceding example, the same



(a)

(b)

Figure 11.9 Root locus for the closed-loop system of eq. (11.40) as the value of β is varied: (a) $\beta > 0$; (b) $\beta < 0$. Note that we have marked the point $z = 2$ that corresponds to the pole location when $\beta = 0$.

will be true for the closed-loop system function

$$Q(s) = \frac{s}{s - 2(1 - \beta)} \tag{11.43}$$

vis-à-vis $Q(z)$, and the locus of the pole as a function of β will be identical to the locus in that example.

The relationship between these two examples stresses the fact that the locus of the poles is determined by the algebraic expressions for the system functions of the forward and feedback paths and is not inherently associated with whether the system is a continuous-time or discrete-time system. However, the interpretation of the result is intimately connected with its continuous-time or discrete-time context. In the discrete-time case it is the location of the poles in relation to the unit circle that is important, whereas in the continuous-time case it is their location in relation to the imaginary axis. Thus, as we have seen for the discrete-time example in eq. (11.40), the system is stable for $1/2 < \beta < 3/2$, while the continuous-time system of eq. (11.43) is stable for $\beta > 1$.

11.3.2 Equation for the Closed-Loop Poles

In the simple example considered in the previous section the root locus was easy to plot, since we could first explicitly determine the closed-loop pole as a function of the gain parameter and then plot the location of the pole as we changed the gain. For more complex systems, one cannot expect to find such simple closed-form expressions for the closed-loop poles. However, it is still possible to sketch accurately the locus of the poles as the value of the gain parameter is varied from $-\infty$ to $+\infty$, without actually solving for the location of the poles for any specific value of the gain. This technique for determining the root locus is extremely useful in gaining insight into the characteristics of a feedback system. Also, as we develop the method, we will see that once we have determined the root locus, there is a relatively straightforward procedure for determining the value of the gain parameter that produces a closed-loop pole at any specified location along the root locus. We will phrase our discussion in terms of the Laplace transform variable s, with the understanding that it applies equally well to the discrete-time case.

Consider a modification of the basic feedback system of Figure 11.3(a), where either $G(s)$ or $H(s)$ is cascaded with an adjustable gain K. This is illustrated in Figure 11.10. In either of these cases, the denominator of the closed-loop system function is $1 + KG(s)H(s)$.[2] Therefore, the equation for the poles of the closed-loop system are the solutions of the equation

$$1 + KG(s)H(s) = 0. \tag{11.44}$$

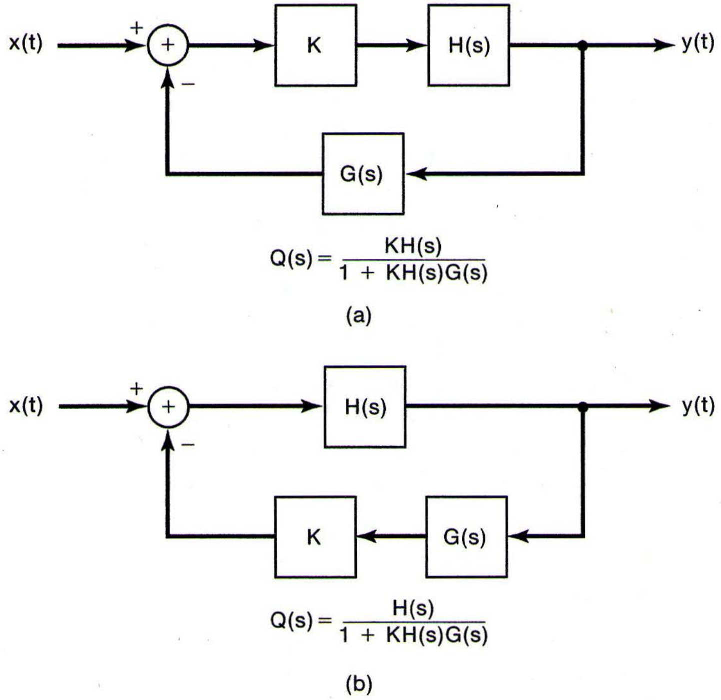

x(t)
+
+
−
K
H(s)
y(t)
G(s)
$Q(s) = \frac{KH(s)}{1 + KH(s)G(s)}$
(a)
x(t)
+
+
−
H(s)
y(t)
K
G(s)
$Q(s) = \frac{H(s)}{1 + KH(s)G(s)}$
(b)

Figure 11.10 Feedback systems containing an adjustable gain: (a) system in which the gain is located in the forward path; (b) system with the gain in the feedback path.

---

[2] In the following discussion, we assume for simplicity that there is no pole-zero cancellation in the product $G(s)H(s)$. The presence of such pole-zero cancellations does not cause any real difficulties, and the procedure we will outline here is easily extended to that case (Problem 11.32). In fact, the simple example at the start of this section [eqs. (11.41) and (11.42)] *does* involve a pole-zero cancellation, at $s = 0$.

Rewriting eq. (11.44), we obtain the basic equation determining the closed-loop poles:

$$G(s)H(s) = -\frac{1}{K}. \tag{11.45}$$

The technique for plotting the root locus is based on the properties of this equation and its solutions. In the remainder of this section, we will discuss some of these properties and indicate how they can be exploited in determining the root locus.

11.3.3 The End Points of the Root Locus: The Closed-Loop Poles for K = 0 and |K| = +∞

Perhaps the most immediate observation that one can make about the root locus is that obtained by examining eq. (11.45) for $K = 0$ and $|K| = \infty$. For $K = 0$, the solution of this equation must yield the poles of $G(s)H(s)$, since $1/K = \infty$. To illustrate, recall the example given by eqs. (11.41) and (11.42). If we let β play the role of K, we see that eq. (11.45) becomes

$$\frac{2}{s-2} = -\frac{1}{\beta}. \tag{11.46}$$

Therefore, for $\beta = 0$, the pole of the system will be located at the pole of $2/(s-2)$ (i.e., at $s = 2$), which agrees with what we depicted in Figure 11.9.

Suppose now that $|K| = \infty$. Then $1/K = 0$, so that the solutions of eq. (11.45) must approach the *zeros* of $G(s)H(s)$. If the order of the numerator of $G(s)H(s)$ is smaller than that of the denominator, then some of these zeros, equal in number to the difference in order between the denominator and numerator, will be at infinity.

Referring again to eq. (11.46), since the order of the denominator of $2/(s-2)$ is 1, while the order of the numerator is zero, we conclude that in this example there is one zero at infinity and no zeros in the finite s-plane. Thus, as $|\beta| \to \infty$, the closed-loop pole approaches infinity. Again, this agrees with Figure 11.9, in which the magnitude of the pole increases without bound as $|\beta| \to \infty$ for either $\beta > 0$ or $\beta < 0$.

While the foregoing observations provide us with basic information as to the closed-loop pole locations for the extreme values of K, the following result is the key to our being able to plot the root locus without actually solving for the closed-loop poles as explicit functions of the gain.

11.3.4 The Angle Criterion

Consider again eq. (11.45). Since the right-hand side of this equation is real, a point s_0 can be a closed-loop pole only if the left-hand side of the equation, i.e., $G(s_0)H(s_0)$, is also real. Writing

$$G(s_0)H(s_0) = \left|G(s_0)H(s_0)\right| e^{j\measuredangle G(s_0)H(s_0)}, \tag{11.47}$$

we see that, for $G(s_0)H(s_0)$ to be real, it must be true that

$$e^{j\measuredangle G(s_0)H(s_0)} = \pm 1. \tag{11.48}$$

That is, for s_0 to be a closed-loop pole, we must have

$$\sphericalangle G(s_0)H(s_0) = \text{integer multiple of } \pi. \tag{11.49}$$

Returning to eq. (11.46), we see immediately that in order for $2/(s_0 - 2)$ to be real, it is necessary that s_0 be real. For more complex system functions, it is not as easy to determine the values of s_0 for which $G(s_0)H(s_0)$ is real. However, as we will see, the use of the angle criterion given by eq. (11.49), together with the geometric method described in Chapter 9 for evaluating $\sphericalangle G(s_0)H(s_0)$, greatly facilitates the determination of the root locus.

The angle criterion given by eq. (11.49) provides us with a direct method for determining whether a point s_0 could be a closed-loop pole for *some* value of the gain K. A further examination of eq. (11.45) gives us a way in which to calculate the value of the gain corresponding to any point on the root locus. Specifically, suppose that s_0 satisfies

$$\sphericalangle G(s_0)H(s_0) = \text{odd multiple of } \pi. \tag{11.50}$$

Then $e^{j\sphericalangle G(s_0)H(s_0)} = -1$, and from eq. (11.47) we see that

$$G(s_0)H(s_0) = -\left|G(s_0)H(s_0)\right|. \tag{11.51}$$

Substituting eq. (11.51) into eq. (11.45), we find that if

$$K = \frac{1}{\left|G(s_0)H(s_0)\right|}, \tag{11.52}$$

then s_0 is a solution of the equation and hence a closed-loop pole.

Similarly, if s_0 satisfies the condition

$$\sphericalangle G(s_0)H(s_0) = \text{even multiple of } \pi, \tag{11.53}$$

then

$$G(s_0)H(s_0) = \left|G(s_0)H(s_0)\right|. \tag{11.54}$$

Thus, if

$$K = -\frac{1}{\left|G(s_0)H(s_0)\right|}, \tag{11.55}$$

then s_0 is a solution of eq. (11.45) and hence a closed-loop pole.

For the example given in eq. (11.46), if s_0 is on the real line and $s_0 < 2$, then

$$\sphericalangle\left(\frac{2}{s_0 - 2}\right) = -\pi, \tag{11.56}$$

and from eq. (11.52), the value of β for which s_0 is the closed-loop pole is

$$\beta = \frac{1}{\left|\frac{2}{s_0-2}\right|} = \frac{2 - s_0}{2}. \tag{11.57}$$

That is,

$$s_0 = 2(1 - \beta), \tag{11.58}$$

which agrees with eq. (11.43).

Summarizing the last two observations that we have made, we see that the *root locus* for the closed-loop system, that is, the set of points in the complex s-plane that are closed-loop poles for *some* value of K as K varies from $-\infty$ to $+\infty$, are precisely those points that satisfy the angle condition of eq. (11.49). Furthermore:

1. A point s_0 for which

$$\sphericalangle G(s_0)H(s_0) = \text{odd multiple of } \pi \tag{11.59}$$

is on the root locus and is a closed-loop pole for some value of $K > 0$. The value of the gain that makes s_0 a closed-loop pole is given by eq. (11.52).

2. A point s_0 for which

$$\sphericalangle G(s_0)H(s_0) = \text{even multiple of } \pi \tag{11.60}$$

is on the root locus and is a closed-loop pole for some value of $K < 0$. The value of the gain that makes s_0 a closed-loop pole is given by eq. (11.55).

Therefore, we have now reduced the problem of determining the root locus to that of searching for points that satisfy the angle requirements given by eqs. (11.59) and (11.60). These equations can be refined further to a set of properties that aid in sketching the root locus. Before discussing these properties, however, let us consider a simple example.

Example 11.1

Let

$$H(s) = \frac{1}{s+1}, \qquad G(s) = \frac{1}{s+2}. \tag{11.61}$$

Recall that in Section 9.4 we discussed the geometric evaluation of Laplace transforms. In that section, we saw that the angle of the rational Laplace transform

$$\frac{\prod_{k=1}^{m}(s-\beta_k)}{\prod_{k=1}^{n}(s-\alpha_k)} \tag{11.62}$$

evaluated at some point s_0 in the complex plane equals the sum of the angles of the vectors from each of the zeros to s_0 minus the sum of the angles from each of the poles to s_0. Applying this to the product of $G(s)H(s)$, where $G(s)$ and $H(s)$ are as given in eq. (11.61), we can determine geometrically those points in the s-plane that satisfy the angle criteria, eqs. (11.59) and (11.60), and therefore can sketch the root locus.

In Figure 11.11, we have plotted the poles of $G(s)H(s)$ and have denoted by θ and ϕ the angles from each of the poles to the point s_0. Let us first test the angle criterion

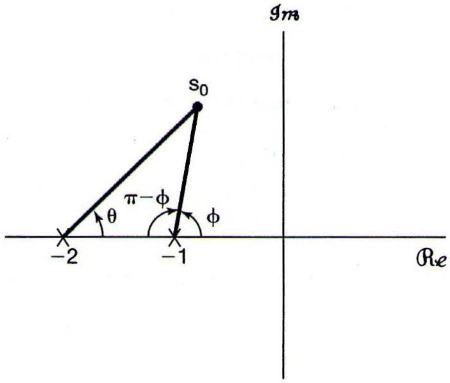


Figure 11.11 Geometric procedure for evaluating angle criterion in Example 11.1.

for points s_0 on the real axis. To begin with, the angle contribution from both poles is zero when s_0 is on the real axis to the right of -1. Thus,

$$\sphericalangle G(s_0)H(s_0) = 0 = 0 \cdot \pi, \qquad s_0 \text{ real and greater than } -1, \tag{11.63}$$

and by eq. (11.60), these points are on the root locus for $K < 0$. For points between the two poles, the pole at -1 contributes an angle of $-\pi$, and the pole at -2 contributes 0. Thus,

$$\sphericalangle G(s_0)H(s_0) = -\pi, \qquad s_0 \text{ real}, -2 < s_0 < -1. \tag{11.64}$$

These points are on the locus for $K > 0$. Finally, each pole contributes an angle of $-\pi$ when s_0 is real and less than -2, so that

$$\sphericalangle G(s_0)H(s_0) = -2\pi, \qquad s_0 \text{ real and less than } -2.$$

Therefore, these points are on the locus for $K < 0$.

Let us now examine points in the upper half of the s-plane. (Since the impulse responses are real-valued, the complex poles occur in conjugate pairs. Therefore, we can immediately determine the poles in the lower half-plane after we have examined the upper half.) From Figure 11.11, the angle of $G(s_0)H(s_0)$ at the point s_0 is

$$\sphericalangle G(s_0)H(s_0) = -(\theta + \phi). \tag{11.65}$$

Also, it is clear that as s_0 ranges over the upper half-plane (but not the real axis), we have

$$0 < \theta < \pi, \qquad 0 < \phi < \pi. \tag{11.66}$$

Thus,

$$-2\pi < \sphericalangle G(s)H(s) < 0, \tag{11.67}$$

Therefore, we see immediately that *no* point in the upper half-plane can be on the locus for $K < 0$ [since $\sphericalangle G(s)H(s)$ never equals an even multiple of π]. In addition, if s_0 is to be on the locus for $K > 0$, we must have

$$\sphericalangle G(s_0)H(s_0) = -(\theta + \phi) = -\pi, \tag{11.68}$$

or

$$\theta = \pi - \phi. \tag{11.69}$$

Examining the geometry of Figure 11.11, we see that this occurs only for those points located on the straight line that is parallel to the imaginary axis and that bisects the line joining the poles at -1 and -2. We have now examined the entire s-plane and have determined all those points on the root locus. In addition, we know that for $K = 0$, the closed-loop poles equal the poles of $G(s)H(s)$, and as $|K| \to \infty$, the closed-loop poles go to the zeros of $G(s)H(s)$, which in this case are both at infinity. Putting these results together, we can draw the entire root locus, depicted in Figure 11.12, in which we have indicated the direction of increasing $|K|$, both for $K > 0$ and for $K < 0$.

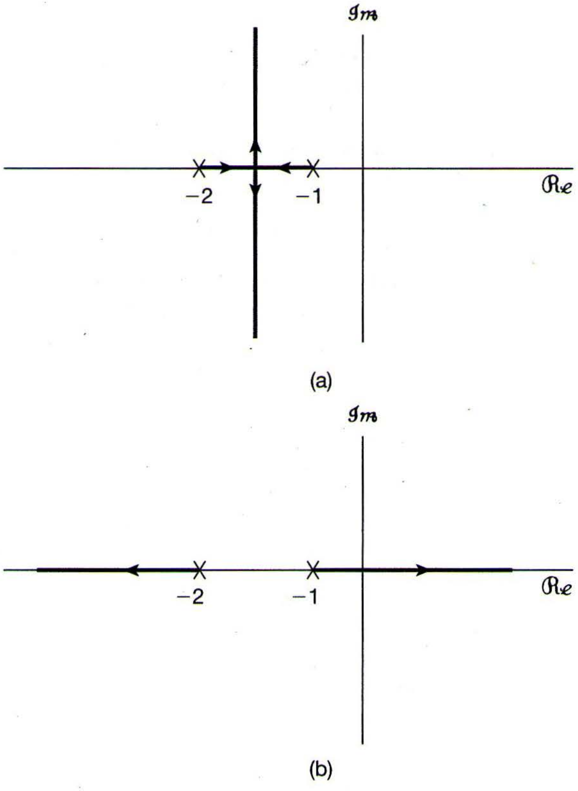


Figure 11.12 Root locus for Example 11.1: (a) $K > 0$; (b) $K < 0$. The poles of $G(s)H(s)$, which are located at $s = -1$ and $s = -2$, are indicated.

Note from the figure that for $K > 0$ there are two branches of the root locus and that the same is true for $K < 0$. The reason for the existence of two branches is that in this example the closed-loop system is a second-order system and consequently has two poles for any specified value of K. Therefore, the root locus has two branches, each of which traces the location of one of the closed-loop poles as K is varied, and for any particular value of K, there is one closed-loop pole on each branch. Again, if we wish to calculate the value of K for which a specific point s_0 on the locus is a closed-loop pole, we can use eqs. (11.52) and (11.55).

11.3.5 Properties of the Root Locus

The procedure outlined in the preceding subsection provides us, in principle, with a method for determining the root locus for any continuous-time or discrete-time LTI feedback system. That is, we simply determine, graphically or otherwise, all those points that satisfy eq. (11.59) or eq. (11.60). Fortunately, there are a number of other geometric properties concerning root loci that make the sketching of a locus far less tedious. To begin our discussion of these properties, let us assume that we have placed $G(s)H(s)$ in the standard form

$$G(s)H(s) = \frac{s^m + b_{m-1}s^{m-1} + \cdots + b_0}{s^n + a_{n-1}s^{n-1} + \cdots + a_0} = \frac{\prod_{k=1}^{m}(s - \beta_k)}{\prod_{k=1}^{n}(s - \alpha_k)}, \tag{11.70}$$

where the β_k's denote the zeros and the α_k's denote the poles. In general, these may be complex. Also, from eq. (11.70), we see that we are assuming that the leading coefficient in both the numerator and the denominator of $G(s)H(s)$ is $+1$. This can always be achieved by dividing the numerator and denominator by the denominator coefficient of s^n and absorbing the resulting numerator coefficient of s^m into the gain K. For example,

$$K\frac{2s+1}{3s^2+5s+2} = K\frac{\frac{2}{3}s+\frac{1}{3}}{s^2+\frac{5}{3}s+\frac{2}{3}} = \left(\frac{2}{3}\right)K\frac{s+\frac{1}{2}}{s^2+\frac{5}{3}s+\frac{2}{3}}, \tag{11.71}$$

and the quantity (2/3)K is then regarded as the overall gain that is varied in determining the root locus.

We assume, in addition, that

$$m \leq n, \tag{11.72}$$

which is the case that is usually encountered in practice. (Problem 11.33 considers the case $m > n$.) The following are some properties that include earlier observations and that aid in sketching the root locus.

Property 1: For $K = 0$, the solutions of eq. (11.45) are the poles of $G(s)H(s)$. Since we are assuming n poles, the root locus has n branches, each one starting (for $K = 0$) at a pole of $G(s)H(s)$.

Property 1 includes the general version of a fact which we noted in Example 11.1: that there is one branch of the root locus for each closed-loop pole. The next property is also simply a restatement of one of our earlier observations.

Property 2: As $|K| \to \infty$, each branch of the root locus approaches a zero of $G(s)H(s)$. Since we are assuming that $m \leq n$, $n - m$ of these zeros are at infinity.

Property 3: Parts of the real s-axis that lie to the left of an *odd* number of real poles and zeros of $G(s)H(s)$ are on the root locus for $K > 0$. Parts of the real s-axis that lie to the left of an *even* number (possibly zero) of poles and zeros of $G(s)H(s)$ are on the root locus for $K < 0$.

We can show that Property 3 is true as follows: From our discussion in Example 11.1 and from Figure 11.13(a), we see that if a point on the real s-axis is to the right of a real pole or zero of $G(s)H(s)$, that pole or zero contributes zero to $\sphericalangle G(s_0)H(s_0)$. On the other hand, if s_0 is to the left of a zero, that zero contributes $+\pi$, whereas if s_0 is to the left of a pole, we get a contribution of $-\pi$ (since we subtract the pole angles). Hence, if s_0 is to the left of an odd number of real poles and zeros, the total contribution of these poles and zeros is an odd multiple of π, whereas if s_0 is to the left of an even number of real poles and zeros, the total contribution is an even multiple of π. From eqs. (11.59) and (11.60), we will have the result stated in Property 3 if we can show that the total contribution from all poles and zeros with nonzero imaginary parts is an even multiple of π. The key here is that such poles and zeros occur in complex-conjugate pairs, and we can consider the contribution from each such pair, as illustrated in Figure 11.13(b). The symmetry in the picture clearly indicates that the sum of the angles from this pair to any point s_0 on the real axis is precisely 2π. Summing over all conjugate zero pairs and subtracting the sum over all conjugate pole pairs, we get the desired result. Thus, *any* segment of the real line

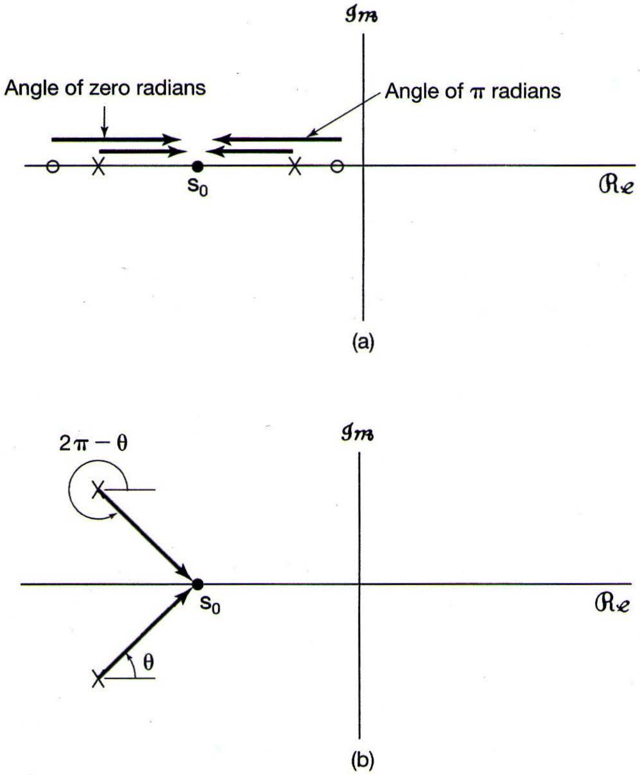


Figure 11.13 (a) Angle contribution from real poles and zeros to a point on the real axis; (b) total angle contribution from a complex-conjugate pole pair to a point on the real axis.

between real poles or zeros is on the root locus either for $K > 0$ or for $K < 0$, depending on whether it lies to the left of an odd or an even number of poles and zeros of $G(s)H(s)$.

As one consequence of Properties 1 through 3, consider a segment of the real axis between two poles of $G(s)H(s)$, with no zeros between these poles. From Property 1 the root locus begins at the poles, and from Property 3 the entire portion of the real axis between the two poles will lie on the root locus for a positive or negative range of values of K. Therefore, as $|K|$ increases from zero, the two branches of the root locus that begin at these poles move toward each other along the segment of the real axis between the poles. From Property 2, as $|K|$ increases toward infinity, each branch of the root locus must approach a zero. Since there are no zeros along that portion of the real axis, the only way that this can happen is if the branches break off into the complex plane for $|K|$ sufficiently large. This is illustrated in Figure 11.12, in which the locus for $K > 0$ has a portion between two real poles. As K is increased, the root locus eventually leaves the real axis, forming two complex-conjugate branches. Summarizing this discussion, we have the following property of the root locus:

Property 4: Branches of the root locus between two real poles must break off into the complex plane for $|K|$ large enough.

Properties 1–4 serve to illustrate how characteristics of the root locus can be deduced from eqs. (11.45), (11.59), and (11.60). In many cases, plotting the poles and zeros of $G(s)H(s)$ and then using these four properties suffices to provide a reasonably accurate sketch of the root locus. (See Examples 11.2 and 11.3.) In addition to these properties, however, there are numerous other characteristics of the root locus that allow one to obtain sketches of increasing accuracy. For example, from Property 2, we know that $n - m$ branches of the root locus approach infinity. In fact, these branches approach infinity at specific angles that can be calculated, and therefore, the branches are asymptotically parallel to lines at these angles. Moreover, it is possible to draw in the asymptotes and then determine the point at which they intersect. These two properties and several others are illustrated in Problems 11.34–11.36 and 11.41–11.42. A more detailed development of the root-locus method can be found in more advanced texts such as those listed in the bibliography at the end of the book.

In the remainder of this section we present two examples, one in continuous time and one in discrete time, that illustrate how the four properties that we have just described allow us to sketch the root locus and to deduce the stability characteristics of a feedback system as the gain K is varied.

Example 11.2

Let

$$G(s)H(s) = \frac{s-1}{(s+1)(s+2)}. \tag{11.73}$$

From Properties 1 and 2, the root locus for either K positive or K negative starts at the points $s = -1$ and $s = -2$. One branch terminates at the zero at $s = 1$ and the other at infinity.

Let us first consider $K > 0$. The root locus in this case is illustrated in Figure 11.14(a). From Property 3, we can identify the regions of the real axis that are on

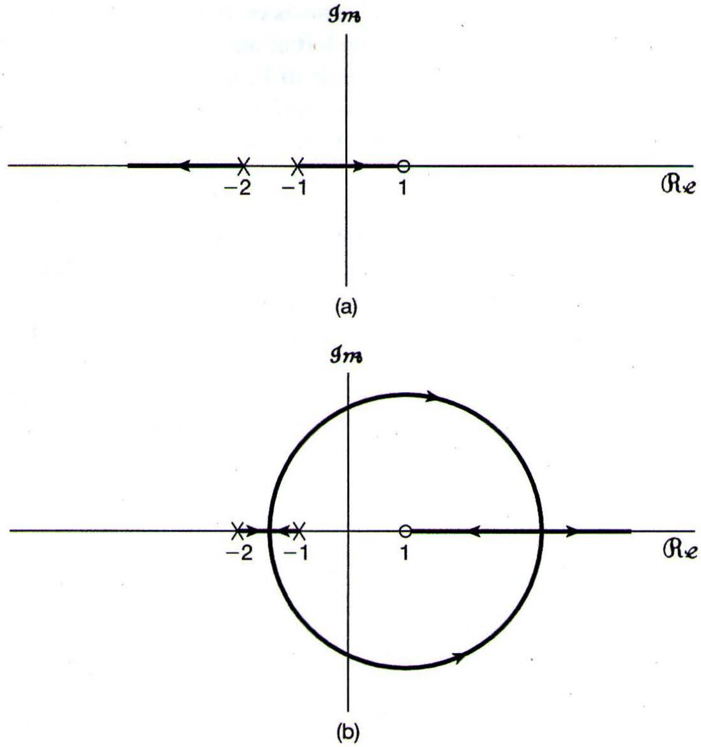


Figure 11.14 Root locus for Example 11.2: (a) $K > 0$; (b) $K < 0$. The poles of $G(s)H(s)$ at $s = -1$ and $s = -2$ and the zero of $G(s)H(s)$ at $s = 1$ are indicated in the figure.

the root locus for $K > 0$—specifically, $\Re e\{s\} < -2$ and $-1 < \Re e\{s\} < 1$. Therefore, one branch of the root locus for $K > 0$ originates at $s = -1$ and approaches $s = 1$ as $K \to \infty$. The other begins at $s = -2$ and extends to the left toward $\Re e\{s\} = -\infty$ as $K \to +\infty$.

Thus, we see that for $K > 0$, if K is sufficiently large, the system will become unstable, as one of the closed-loop poles moves into the right-half plane. The procedure that we have used for sketching the root locus does not, of course, indicate the value of K for which this instability develops. However, for this particular example, we see that the value of K for which the instability occurs corresponds to the root locus passing through $s = 0$. Consequently, from eq. (11.52), the corresponding value of K is

$$K = \frac{1}{|G(0)H(0)|} = 2. \tag{11.74}$$

Thus, the system is stable for $0 \leq K < 2$, but is unstable for $K \geq 2$.

For $K < 0$, the portions of the real axis lying on the root locus are $\Re e\{s\} > 1$ and $-2 < \Re e\{s\} < -1$. Thus, the root locus again starts at the points $s = -2$ and $s = -1$, moving into the region $-2 < \Re e\{s\} < -1$. At some point, it breaks off into the complex plane and follows a trajectory such that it returns to the real axis for $s > 1$. Upon the root locus' return to the real axis, one branch moves to the left toward the zero at $s = 1$ and the other to the right toward $s = \infty$, as indicated in Figure 11.14(b), in which we have displayed an accurate plot of the root locus for $K < 0$.

Rules can also be developed to indicate the locations at which the root locus leaves and enters the real axis. Even without that precise a description, however, we can sketch the general shape of the root locus in Figure 11.14(b) and can therefore deduce that for $K < 0$, the system also becomes unstable for $|K|$ sufficiently large.

Example 11.3

Consider the discrete-time feedback system illustrated in Figure 11.15. In this case,

$$G(z)H(z) = \frac{z^{-1}}{\left(1 - \frac{1}{2}z^{-1}\right)\left(1 - \frac{1}{4}z^{-1}\right)} = \frac{z}{\left(z - \frac{1}{2}\right)\left(z - \frac{1}{4}\right)}. \tag{11.75}$$

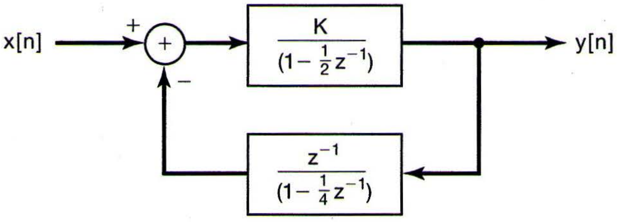

x[n]
+
−
$\frac{K}{(1-\frac{1}{2}z^{-1})}$
$\frac{z^{-1}}{(1-\frac{1}{4}z^{-1})}$
y[n]

Figure 11.15 Discrete-time feedback system of Example 11.3.

As discussed at the beginning of this section, the techniques for sketching the root locus of a discrete-time feedback system are identical to those used in the continuous-time case. Therefore, in a manner exactly analogous to that of the preceding example, we can deduce the basic form of the root locus for this example, which is illustrated in Figure 11.16. In this case the portion of the real axis between the two poles of $G(z)H(z)$ (at $z = 1/4$ and $z = 1/2$) is on the root locus for $K > 0$, and as K increases the locus breaks off into the complex plane and returns to the axis at some point in the left-half plane. From there, one branch approaches the zero of $G(z)H(z)$ at $z = 0$, and the other approaches infinity as $K \to \infty$. The form of the root locus for $K < 0$ consists of two branches on the real axis, one approaching 0 and the other infinity.

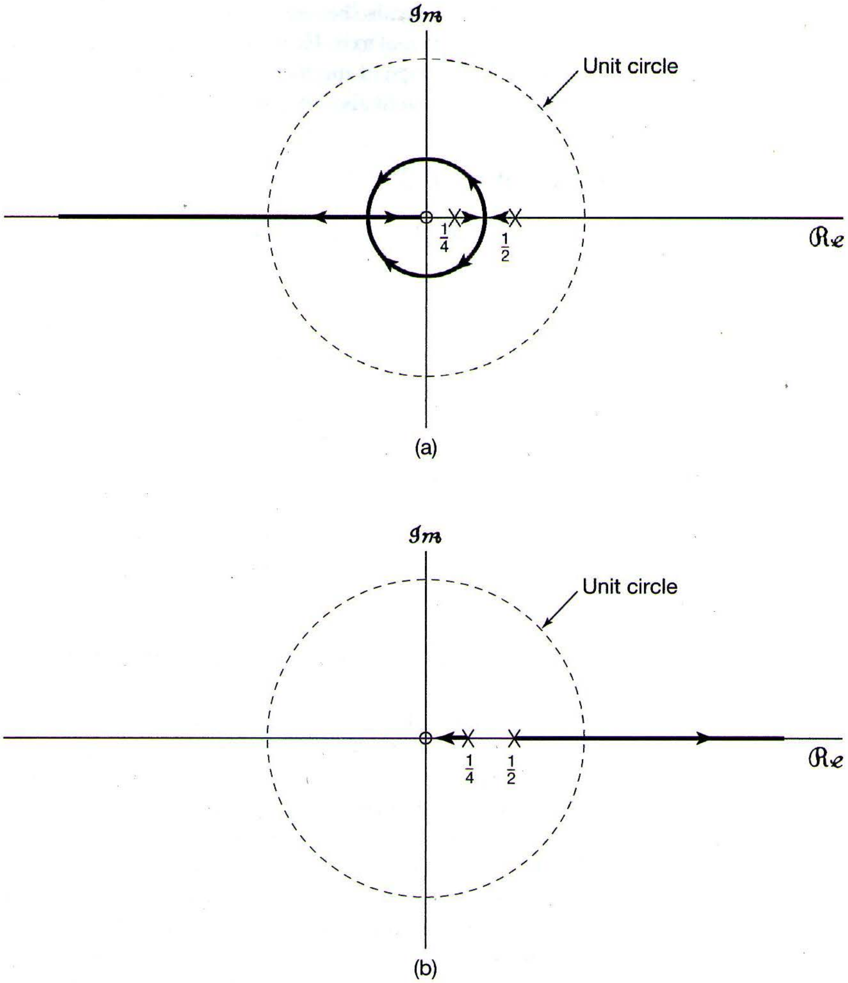
As we remarked earlier, while the form of the root locus does not depend on whether the system is a continuous-time or discrete-time system, any conclusion regarding stability based on examining the locus certainly does. For this example, we can conclude that for $|K|$ sufficiently large the system is unstable, since one of the two poles has magnitude greater than 1. In particular, from the $K > 0$ root locus in Figure 11.16(a), we see that the transition from stability to instability occurs when one of the closed-loop poles is at $z = -1$. From eq. (11.52), the corresponding value of K is

$$K = \frac{1}{|G(-1)H(-1)|} = \frac{15}{8}. \tag{11.76}$$

Similarly, from Figure 11.16(b), the stability-instability transition occurs when one of the closed-loop poles is at $z = 1$, and from eq. (11.55), the corresponding value of K is

$$K = -\frac{1}{|G(1)H(1)|} = -\frac{3}{8}. \tag{11.77}$$

Putting this together, we see that the closed-loop system in Figure 11.16 is stable if



$\mathcal{I}m$
Unit circle
$\frac{1}{4}$
$\frac{1}{2}$
$\mathcal{R}e$
(a)
$\mathcal{I}m$
Unit circle
$\frac{1}{4}$
$\frac{1}{2}$
$\mathcal{R}e$
(b)

Figure 11.16 Root locus for Example 11.3: (a) $K > 0$; (b) $K < 0$. The poles of $G(z)H(z)$ at $z = 1/4$ and $z = 1/2$ and the zero of $G(z)H(z)$ at $z = 0$ are indicated in the figure.

$$-\frac{3}{8} < K < \frac{15}{8} \tag{11.78}$$

and is unstable for K outside this range.

11.4 THE NYQUIST STABILITY CRITERION

As developed in Section 11.3, the root-locus technique provides detailed information concerning the location of closed-loop poles as the system gain is varied. From root-locus plots, one can determine the damping and the stability characteristics of the system as K is varied. Determination of the root locus requires the analytic description of the system functions of the forward and feedback paths and is applicable only when these transforms are rational. For example, it cannot be directly applied in situations in which our knowledge of the system functions is obtained purely from experimentation.

In this section, we introduce another method for the determination of the stability of feedback systems as a function of an adjustable gain parameter. This technique, referred to as the *Nyquist criterion,* differs from the root-locus method in two basic ways. Unlike the root-locus method, the Nyquist criterion does *not* provide detailed information concerning the location of the closed-loop poles as a function of K, but rather, simply determines whether or not the system is stable for any specified value of K. On the other hand, the Nyquist criterion can be applied to nonrational system functions and in situations in which no analytic description of the forward and feedback path system functions is available.

Our objective in this section is to outline the basic ideas behind the Nyquist criterion for both continuous-time and discrete-time systems. As we will see, both the discrete-time and continuous-time Nyquist tests are the result of the same fundamental concept, although, as with the root-locus method, the actual criteria for stability differ because of the differences between continuous and discrete time. More detailed developments of the ideas behind the Nyquist criterion and its use in the design of feedback systems can be found in texts on the analysis and synthesis of feedback systems and automatic control systems, including those listed in the bibliography at the end of the book.

To introduce the method, let us recall that the poles of the closed-loop systems of Figure 11.10 and their discrete-time counterparts are the solutions of the equations

$$1 + KG(s)H(s) = 0 \qquad \text{(continuous time)} \tag{11.79}$$

and

$$1 + KG(z)H(z) = 0 \qquad \text{(discrete time).} \tag{11.80}$$

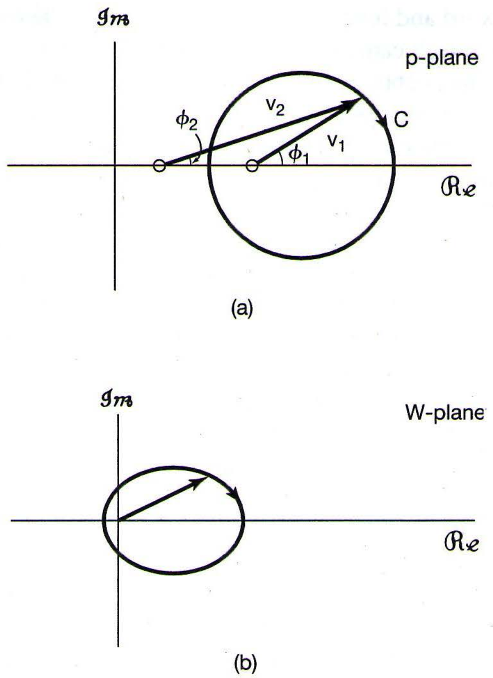
For discrete-time systems, we want to determine whether any of the solutions of eq. (11.80) lie outside the unit circle, and for continuous-time systems whether any of the solutions of eq. (11.79) lie in the right half of the s-plane. The Nyquist criterion fixes this by examination of the values of $G(s)H(s)$ along the $j\omega$-axis and the values of $G(z)H(z)$ along the unit circle. The basis for this is the encirclement property, which we develop in the following subsection.

11.4.1 The Encirclement Property

Consider a general rational function $W(p)$, where p is a complex variable,[3] and suppose that we plot $W(p)$ for values of p along a closed contour in the p-plane, which we traverse in a clockwise direction. This is illustrated in Figure 11.17 for a function $W(p)$ that has two zeros and no poles. In Figure 11.17(a) we show a closed contour C in the p-plane, and in Figure 11.17(b) we plot the closed contour of the values of $W(p)$ as p varies around C.

---

[3] Because we will use the property we are about to develop for both continuous-time and discrete-time feedback systems, we have chosen to describe it in terms of a general complex variable p. In the next subsection we use this property to analyze continuous-time feedback systems, where the complex variable is s. Following this, in Section 11.4.3 we use the encirclement property for discrete-time feedback systems, in which context the complex variable is z.



(a)

(b)

Figure 11.17 Basic encirclement property. The closed curve in (b) represents a plot of the values of $W(p)$ for values of p along the curve C in (a). Here, the arrow on the curve C in (a) indicates the direction in which C is traversed, and the arrow in (b) indicates the corresponding direction along the contour of values of $W(p)$.

In this example, there is one zero of $W(p)$ inside the contour and one zero of $W(p)$ outside the contour. At any point p on C, the angle of $W(p)$ is the sum of the angles of the two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ to the point p. As we traverse the contour once, the angle ϕ_1 of the vector from the zero *inside* the contour encounters a net change of -2π radians, whereas the angle ϕ_2 of the vector from the zero *outside* the contour encounters no net change. Thus, on the plot of $W(p)$, there is a net change in angle of -2π. Said another way, the plot of $W(p)$ in Figure 11.17(b) encircles the origin once in the clockwise direction. More generally, for an arbitrary rational $W(p)$, as we traverse a closed contour in the clockwise direction, any poles and zeros of $W(p)$ outside the contour will contribute no net change to the angle of $W(p)$, whereas each zero inside the contour will contribute a net change of -2π and each pole inside will contribute a net change of $+2\pi$. Since each net change of -2π in $W(p)$ corresponds to one clockwise encirclement of the origin in the plot of $W(p)$, we can state the following basic *encirclement property:*

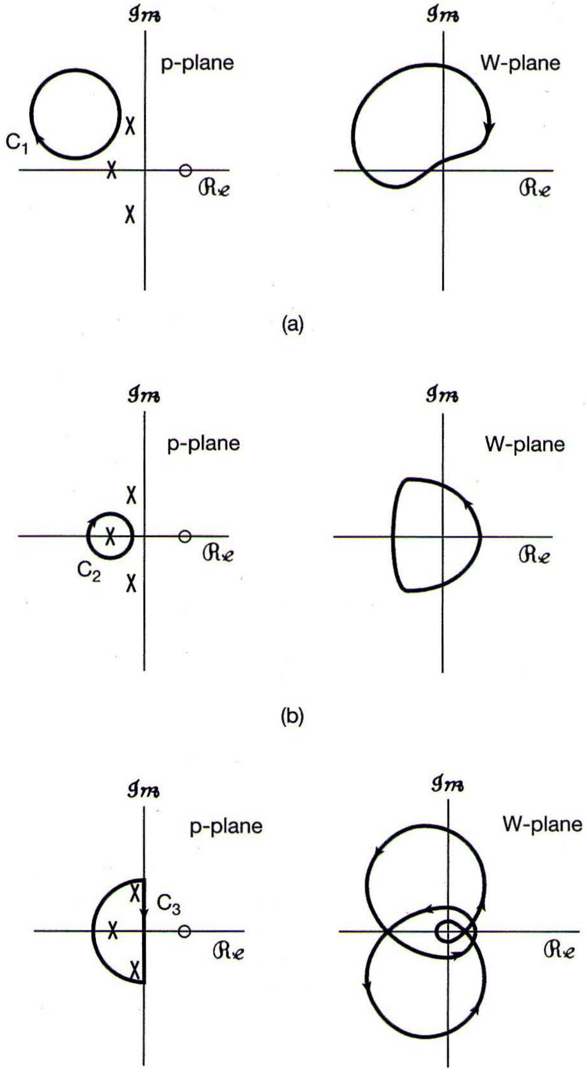
Encirclement Property: As a closed contour C in the p-plane is traversed once in the clockwise direction, the corresponding plot of $W(p)$ for values of p along the contour encircles the origin in the clockwise direction a net number of times equal to the number of zeros minus the number of poles contained within the contour.

In applying this statement, a counterclockwise encirclement is interpreted as the negative of one clockwise encirclement. For example, if there is one pole and no zeros inside the contour, there will be one counterclockwise, or equivalently, minus-one clockwise, encirclement.

Example 11.4

Consider the function

$$W(p) = \frac{p-1}{(p+1)(p^2+p+1)}. \tag{11.81}$$



$\mathfrak{Im}$
p-plane
C_1
X
X
X
$\mathfrak{Re}$
$\mathfrak{Im}$
W-plane
$\mathfrak{Re}$
$\mathfrak{Im}$
p-plane
X
X
X
C_2
$\mathfrak{Re}$
$\mathfrak{Im}$
W-plane
$\mathfrak{Re}$
$\mathfrak{Im}$
p-plane
X
C_3
X
X
$\mathfrak{Re}$
$\mathfrak{Im}$
W-plane
$\mathfrak{Re}$

(a)

(b)

(c)

Figure 11.18 Basic encirclement property for Example 11.4: (a) the contour encircles no poles or zeros and consequently $W(p)$ has no encirclements of the origin; (b) the contour encircles one pole and therefore $W(p)$ has one encirclement of the origin; (c) the contour encircles three poles and therefore $W(p)$ has three encirclements of the origin;

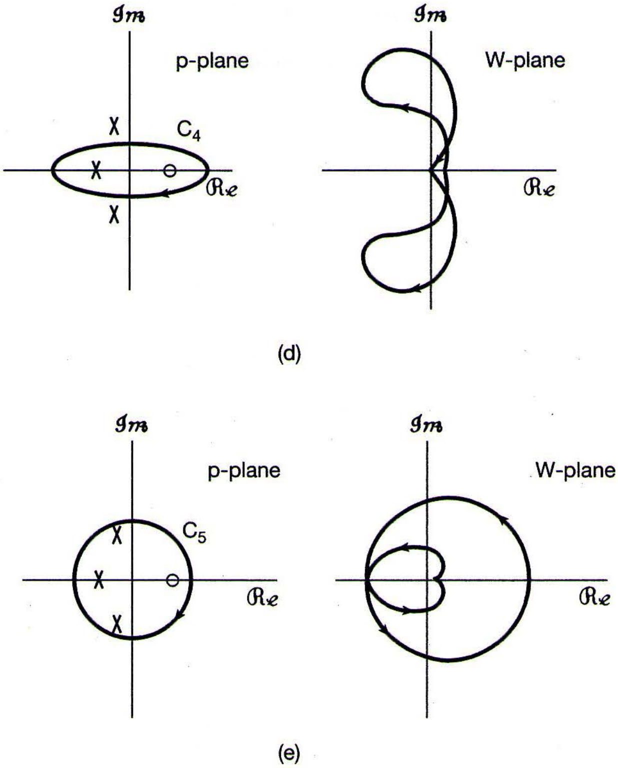


Figure 11.18 *Continued* (d) the contour encircles one pole and one zero and therefore $W(p)$ has no encirclements of the origin; (e) the contour encircles three poles and one zero. $W(p)$ has two encirclements of the origin.

In Figure 11.18, we depict several closed contours in the complex p-plane and the corresponding plots of $W(p)$ along each of these contours. In Figure 11.18(a), the contour C_1 does not encircle any of the poles or zeros of $W(p)$, and consequently, the plot of $W(p)$ has no net encirclements of zero. In Figure 11.18(b), only the pole at $p = -1$ is contained within the contour C_2, and the plot of $W(p)$ encircles the origin once in the counterclockwise direction (minus-one clockwise encirclements). In Figure 11.18(c), C_3 encircles all three poles, and the plot of $W(p)$ encircles the origin three times in a counterclockwise direction. In Figure 11.18(d), C_4 encircles one pole and one zero, and therefore, the plot of $W(p)$ has no net encirclements of the origin. Finally, in Figure 11.18(e), all of the poles and the one zero of $W(p)$ are contained within C_5, and thus, the plot of $W(p)$ along this contour has two net counterclockwise encirclements of the origin.

11.4.2 The Nyquist Criterion for Continuous-Time LTI Feedback Systems

In this section, we exploit the encirclement property in examining the stability of the continuous-time feedback system of Figure 11.10. Stability of this system requires that no zeros of $1 + KG(s)H(s)$, or equivalently, of the function

$$R(s) = \frac{1}{K} + G(s)H(s) \tag{11.82}$$

lie in the right half of the s-plane. Thus, in applying the general result developed in the preceding subsection, we can consider the contour indicated in Figure 11.19. From the plot of $R(s)$, as s traverses the contour C we can obtain a count of the number of zeros minus the number of poles of $R(s)$ contained within the contour by counting the number of clockwise encirclements of the origin. As M increases to infinity, this then corresponds to the number of zeros minus the number of poles of $R(s)$ in the right half of the s-plane.

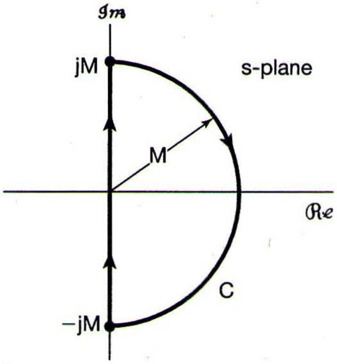


Figure 11.19 Closed contour containing a portion of the right-half plane; as $M \rightarrow \infty$, the contour encloses the entire right-half plane.

Let us examine the evaluation of $R(s)$ along the contour in Figure 11.19 as M increases to infinity. Along the semicircular portion of the contour extending into the right-half plane, we must ensure that $R(s)$ remains bounded as M increases. Accordingly, we will assume that $R(s)$ has at least as many poles as zeros. In that case,

$$R(s) = \frac{b_n s^n + b_{n-1}s^{n-1} + \cdots + b_0}{a_n s^n + a_{n-1}s^{n-1} + \cdots + a_0} \tag{11.83}$$

and

$$\lim_{|s|\rightarrow\infty} R(s) = \frac{b_n}{a_n} = \text{constant.} \tag{11.84}$$

Therefore, as M increases to infinity, the value of $R(s)$ does not change as we traverse the semicircular part of the contour, and consequently, the constant value along this part is equal to the value of $R(s)$ at the end points [i.e., $R(j\omega)$ at $\omega = \pm\infty$].

Therefore, the plot of $R(s)$ along the contour of Figure 11.19 can be obtained by plotting $R(s)$ along the part of the contour that coincides with the imaginary axis—that is, the plot of $R(j\omega)$ as ω varies from $-\infty$ to $+\infty$. Since $R(j\omega)$ equals $1/K + G(j\omega)H(j\omega)$, $R(s)$ along the contour can be drawn from knowledge of $G(j\omega)$ and $H(j\omega)$. If both the forward- and feedback-path systems are stable, $G(j\omega)$ and $H(j\omega)$ are simply the frequency-response functions of these systems. However, the encirclement property for the general function $W(p)$ is simply a property of complex functions; it has nothing to do with whether $W(p)$ arose as the Laplace or z-transform of any signal and, consequently, has nothing to do with regions of convergence. Thus, even if the forward- and feedback-path systems are unstable, if we examine the plot of the *function* $R(j\omega) = 1/K + G(j\omega)H(j\omega)$ for $-\infty < \omega < \infty$, we know that the net number of clockwise encirclements of the origin will equal the number of zeros minus the number of poles of $R(s)$ that lie in the right-half plane.

Furthermore, from eq. (11.82), we see that the poles of $R(s)$ are simply the poles of $G(s)H(s)$, while the zeros of $R(s)$ are the closed-loop poles. In addition, since $G(j\omega)H(j\omega) = R(j\omega) - 1/K$, it follows that the plot of $G(j\omega)H(j\omega)$ encircles the point $-1/K$ *exactly* as many times as $R(j\omega)$ encircles the origin. The plot of $G(j\omega)H(j\omega)$ as ω varies from $-\infty$ to $+\infty$ is called the *Nyquist plot.* From the encirclement property, we then see that

$$\begin{array}{c}\text{The number of clockwise}\\ \text{encirclements of the point}\\ -1/K \text{ by the Nyquist plot}\end{array} = \begin{array}{l}\text{The number of right-half}\\ \text{plane closed-loop poles}\\ \text{minus the number of right-}\\ \text{half plane poles of } G(s)H(s).\end{array} \tag{11.85}$$

While the open-loop system $G(s)H(s)$ may have unstable poles, for the closed-loop system to be stable we require no right-half plane closed-loop poles. This yields the *continuous-time Nyquist stability criterion:*

Continuous-Time Nyquist Stability Criterion: For the closed-loop system to be stable, the net number of *clockwise* encirclements of the point $-1/K$ by the Nyquist plot of $G(j\omega)H(j\omega)$ must equal *minus* the number of right-half-plane poles of $G(s)H(s)$. Equivalently, the net number of *counterclockwise* encirclements of the point $-1/K$ by the Nyquist plot of $G(j\omega)H(j\omega)$ must *equal* the number of right-half-plane poles of $G(s)H(s)$.

For example, if the forward- and feedback-path systems are stable, then the Nyquist plot is simply the plot of the frequency response of the cascade of these two systems. In this case, since there are no poles of $G(s)H(s)$ in the right-half plane, the Nyquist criterion requires that, for stability, the net number of encirclements of the point $-1/K$ must be zero.

Example 11.5

Let

$$G(s) = \frac{1}{s+1}, \qquad H(s) = \frac{1}{\frac{1}{2}s+1}. \tag{11.86}$$

The Bode plot for $G(j\omega)H(j\omega)$ is shown in Figure 11.20. The Nyquist plot depicted in Figure 11.21 is constructed directly from the plots of the log magnitude and phase of $G(j\omega)H(j\omega)$. That is, each point on the Nyquist plot has polar coordinates consisting of the magnitude $|G(j\omega)H(j\omega)|$ and angle $\sphericalangle G(j\omega)H(j\omega)$ for some value of ω. The coordinates of $G(j\omega)H(j\omega)$ for $\omega < 0$ are obtained from the values for $\omega > 0$ through the use of the conjugate symmetry property of $G(j\omega)H(j\omega)$. This property manifests itself geometrically in a very simple way, which facilitates the sketching of the Nyquist plot for any feedback system composed of systems with

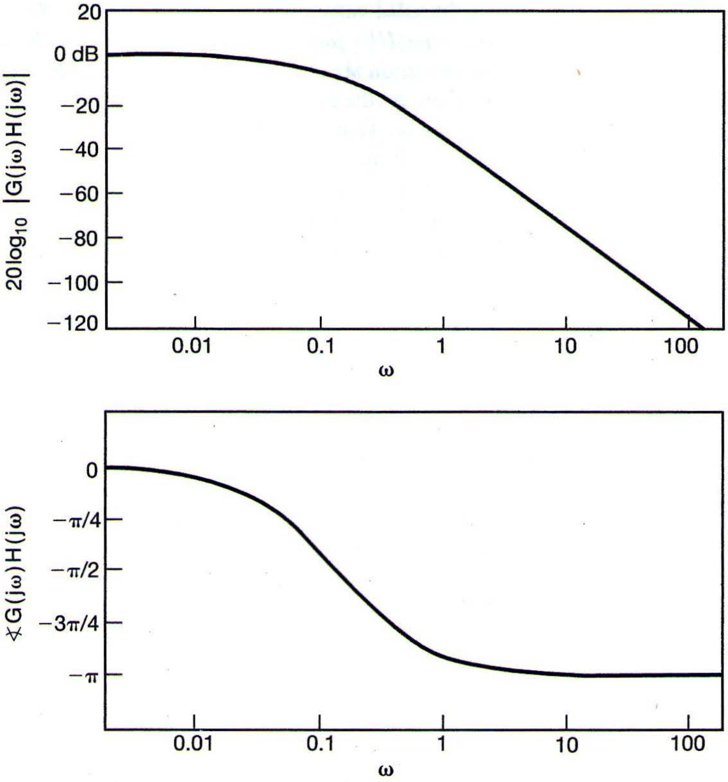


Figure 11.20 Bode plot for $G(j\omega)H(j\omega)$ in Example 11.5.

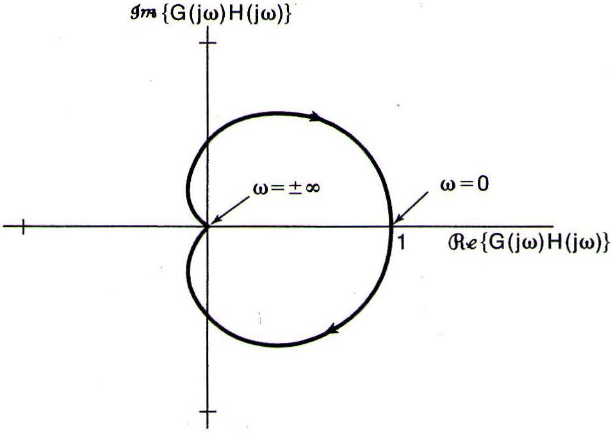


Figure 11.21 Nyquist plot of $G(j\omega)H(j\omega)$ for Example 11.5. The arrow on the curve indicates the direction of increasing ω.

real impulse responses. Specifically, since $|G(-j\omega)H(-j\omega)| = |G(j\omega)H(j\omega)|$ and $\sphericalangle G(-j\omega)H(-j\omega) = -\sphericalangle G(j\omega)H(j\omega)$, the Nyquist plot of $G(j\omega)H(j\omega)$ for $\omega \leq 0$ is a reflection about the real axis of the plot for $\omega \geq 0$. Note also that we have included an arrow on the Nyquist plot in Figure 11.21. This arrow indicates the direction of increasing ω. That is, it indicates the direction in which the Nyquist plot is traversed (as ω varies from $-\infty$ to $+\infty$) for the counting of encirclements in the application of the Nyquist criterion.

In this example there are no right-half-plane open-loop poles, and consequently, the Nyquist criterion requires that, for stability, there be no net encirclements of the point $-1/K$. Thus, by inspection of Figure 11.21, we see that the closed-loop system will be stable if the point $-1/K$ falls outside the Nyquist contour—that is, if

$$-\frac{1}{K} \leq 0 \quad \text{or} \quad -\frac{1}{K} > 1, \tag{11.87}$$

which is equivalent to

$$K \geq 0 \quad \text{or} \quad 0 > K > -1. \tag{11.88}$$

Combining these two conditions, we obtain the result that the closed-loop system will be stable for any choice of K greater than -1.

Example 11.6

Consider now

$$G(s)H(s) = \frac{s+1}{(s-1)(\frac{1}{2}s+1)}. \tag{11.89}$$

The Nyquist plot for this system is indicated in Figure 11.22. For this example, $G(s)H(s)$ has one right-half-plane pole. Thus, for stability we require one counterclockwise encirclement of the point $-1/K$, which in turn requires that the point $-1/K$ fall inside the contour. Hence, we will have stability if and only if $-1 < -1/K < 0$, that is, if $K > 1$.

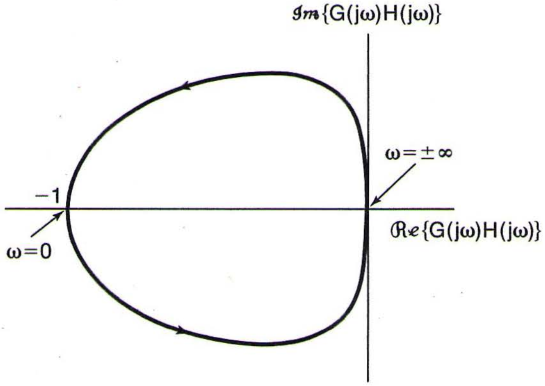


Figure 11.22 Nyquist plot for Example 11.6. The arrow on the curve indicates the direction of increasing ω.

In the foregoing discussion, we have introduced and illustrated a form for the Nyquist stability criterion that applies to an extremely large class of feedback systems. In addition, there are a number of refinements and extensions of the criterion that allow it to be used for many other feedback systems as well. For example, as we have developed it, the Nyquist plot can be drawn without any difficulties for stable or unstable $G(s)H(s)$, as long as there are no poles of $G(s)H(s)$ exactly *on* the $j\omega$-axis. When such poles do occur, the value of $G(j\omega)H(j\omega)$ is infinite at those points. However, as considered in Problem 11.44, the Nyquist criterion can be modified to allow for poles of $G(s)H(s)$ on the $j\omega$-axis. In addition, as mentioned at the beginning of this section, the Nyquist criterion can also be extended to the case in which $G(s)$ and $H(s)$ are not rational. For example, it can be shown that if the forward- and feedback-path systems are both stable, the Nyquist criterion is the same when the system functions are nonrational as it is when they are rational. That is, the closed-loop system is stable if there are no net encirclements of the point $-1/K$. To illustrate the application of the Nyquist criterion for nonrational system functions, we present the following example:

Example 11.7

Consider the acoustic feedback example discussed in Section 11.2.6. Referring to Figure 11.8(b), let $K = K_1K_2$ and

$$G(s)H(s) = -e^{-sT} = e^{-(sT+j\pi)}, \tag{11.90}$$

where we have used the fact that $e^{-j\pi} = -1$. In this case,

$$G(j\omega)H(j\omega) = e^{-j(\omega T+\pi)}, \tag{11.91}$$

and as ω varies from $-\infty$ to ∞, $G(j\omega)H(j\omega)$ traces out a circle of radius 1 in the clockwise direction, with one full revolution for every change of $2\pi/T$ in ω. This is illustrated in Figure 11.23. Since the forward- and feedback-path systems are stable [the cascade $G(s)H(s)$ is simply a time delay], the Nyquist stability criterion indicates that the closed-loop system will be stable if and only if $-1/K$ does not fall inside the unit circle. Equivalently, we require for stability that

$$|K| < 1. \tag{11.92}$$

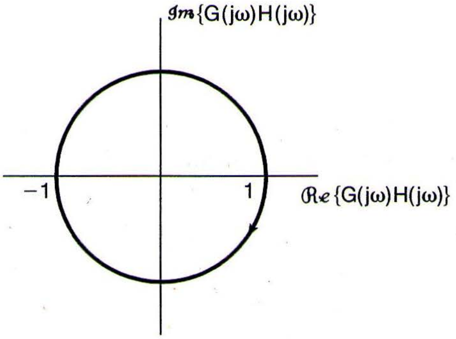

$\mathcal{I}m\{G(j\omega)H(j\omega)\}$
−1
1
$\mathcal{R}e\{G(j\omega)H(j\omega)\}$

Figure 11.23 Nyquist plot for Example 11.7.

Since K_1 and K_2 represent an acoustic gain and attenuation, respectively, they are both positive, which yields the stability criterion

$$K_1K_2 < 1. \tag{11.93}$$

11.4.3 The Nyquist Criterion for Discrete-Time LTI Feedback Systems

As in the continuous-time case, the Nyquist stability criterion for discrete-time systems is based on the fact that the difference in the number of poles and zeros inside a contour, for a rational function, can be determined by examining a plot of the value of the function along the contour. The difference between the continuous-time and discrete-time cases is the choice of the contour. For the discrete-time case, stability of the closed-loop feedback system requires that no zeros of

$$R(z) = \frac{1}{K} + G(z)H(z) \tag{11.94}$$

lie outside the unit circle.

Recall that the encirclement property relates to poles and zeros *inside* any specified contour. On the other hand, in examining the stability of a discrete-time system, we are concerned with the zeros of $R(z)$ *outside* the unit circle. Therefore, in order to make use of the encirclement property, we first make a simple modification. Let us consider the rational function

$$\hat{R}(z) = R\left(\frac{1}{z}\right) \tag{11.95}$$

obtained by replacing z by its reciprocal. As seen in Problem 10.43, if z_0 is a zero (pole) of $R(z)$, then $1/z_0$ is a zero (pole) of $\hat{R}(z)$. Since $1/|z_0|$ is less than 1 if $|z_0| > 1$, any zero or pole of $R(z)$ *outside* the unit circle corresponds to a zero or pole of $\hat{R}(z)$ *inside* the unit circle.

From the basic encirclement property, we know that as z traverses the unit circle in a clockwise direction, the net number of clockwise encirclements of the origin by $\hat{R}(z)$ equals the difference between the number of its zeros and poles inside the unit circle. However, from the previous paragraph, this equals the difference between the number of zeros and poles of $R(z)$ *outside* the unit circle. Furthermore, on the unit circle, $z = e^{j\omega}$ and $1/z = e^{-j\omega}$. Therefore,

$$\hat{R}(e^{j\omega}) = R(e^{-j\omega}). \tag{11.96}$$

From this, we see that evaluating $\hat{R}(z)$ as z traverses the unit circle in the clockwise direction is identical to evaluating $R(z)$ as z traverses the unit circle in the *counterclockwise* direction. In sum, then,

$$\left[\begin{array}{l}\text{The number of clockwise}\\ \text{encirclements of the origin}\\ \text{by the plot of } R(e^{j\omega}) \text{ as the}\\ \text{unit circle is traversed in the}\\ \text{counterclockwise direction}\\ \text{(e.g., as } \omega \text{ increases from 0}\\ \text{to } 2\pi)\end{array}\right] = \left[\begin{array}{l}\text{The number of zeros of } R(z)\\ \text{outside the unit circle minus}\\ \text{the number of poles of } R(z)\\ \text{outside the unit circle.}\end{array}\right] \tag{11.97}$$

Much as in the continuous-time case, counting the encirclements of the origin by $R(e^{j\omega})$ is equivalent to counting the number of encirclements of the point $-1/K$ by the plot of $G(e^{j\omega})H(e^{j\omega})$, again referred to as the Nyquist plot, which is graphed as ω varies from 0 to 2π. Also, the poles of $R(z)$ are precisely the poles of $G(z)H(z)$, and the zeros of $R(z)$ are the closed-loop poles. Therefore, the encirclement property stated in the preceding paragraph implies that the net number of clockwise encirclements by the Nyquist plot of the point $-1/K$ equals the number of closed-loop poles outside the unit circle minus the number of poles of $G(z)H(z)$ outside the unit circle. In order that the closed-loop system be stable, we require no closed-loop poles outside the unit circle. This yields the *discrete-time Nyquist stability criterion*:

Discrete-Time Nyquist Stability Criterion: For the closed-loop system to be stable, the net number of *clockwise* encirclements of the point $-1/K$ by the Nyquist plot of $G(e^{j\omega})H(e^{j\omega})$ as ω varies from 0 to 2π must equal *minus* the number of poles of $G(z)H(z)$ that lie outside the unit circle. Equivalently, the net number of *counterclockwise* encirclements of the point $-1/K$ by the Nyquist plot of $G(e^{j\omega})H(e^{j\omega})$ as ω varies from 0 to 2π must *equal* the number of poles of $G(z)H(z)$ outside the unit circle.

Example 11.8

Let

$$G(z)H(z) = \frac{z^{-2}}{1 + \frac{1}{2}z^{-1}} = \frac{1}{z(z + \frac{1}{2})}. \tag{11.98}$$

The Nyquist plot of this curve is shown in Figure 11.24. Since $G(z)H(z)$ has no poles outside the unit circle, for the stability of the closed-loop system there must be no encirclements of the point $-1/K$. From the figure, we see that this will be the case either if $-1/K < -1$ or if $-1/K > 2$. Thus, the system is stable for $-1/2 < K < 1$.

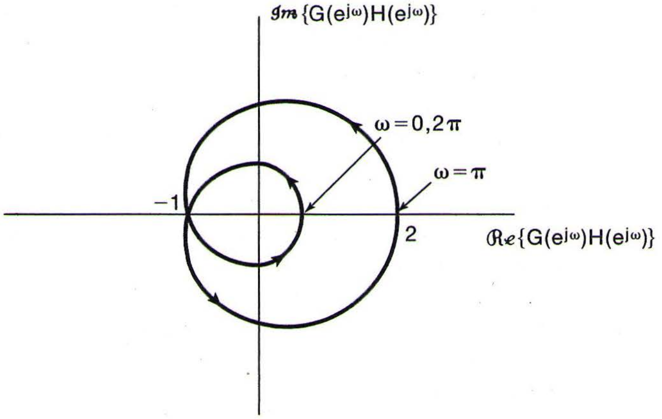


Figure 11.24 Nyquist plot for Example 11.8. The arrow on the curve indicates the direction in which the curve is traversed as ω increases from 0 to 2π.

Just as in continuous time, if the forward and feedback paths are stable, then the Nyquist plot can be obtained from the frequency responses $H(e^{j\omega})$ and $G(e^{j\omega})$ of these systems. If the forward and feedback paths are unstable, then the frequency responses are not defined. Nevertheless, the *function* $G(z)H(z)$ can still be evaluated on the unit circle, and the Nyquist stability criterion can be applied.

As we have seen in this section, the Nyquist stability criterion provides a useful method for determining the range of values of the gain K for which a continuous-time or discrete-time feedback is stable (or unstable). This criterion and the root-locus method are extremely important tools in the design and implementation of feedback systems, and each has its own uses and limitations. For example, the Nyquist criterion can be applied to nonrational system functions, whereas the root-locus method cannot. On the other hand, root-locus plots allow us to examine not only stability, but also other characteristics of the closed-loop system response, such as damping, oscillation frequency, and so on, which are readily identifiable from the location of the poles of the closed-loop system. In the next section, we introduce an additional tool for the analysis of feedback systems that highlights another important characteristic of closed-loop system behavior.

11.5 GAIN AND PHASE MARGINS

In this section, we introduce and examine the concept of the *margin of stability* in a feedback system. It is often of interest not only to know *whether* a feedback system is stable, but also to determine how much the gain in the system can be perturbed and how much additional phase shift can be added to the system before it becomes unstable. Information such as this is important because in many applications the forward and feedback system functions are known only approximately or may change slightly during operation because of wear, the effect of high temperatures on components, or similar influences.

As an example, consider the telescope-pointing system described in Section 11.0 and illustrated in Figures 11.1(c) and (d). This system consists of a motor, a potentiometer converting the shaft angle to a voltage, and an amplifier that is used to amplify the voltage representing the difference between the desired and the actual shaft angles. Assuming that we have obtained approximate descriptions of each of these components, we can set the amplifier gain so that the system will be stable if these descriptions are accurate. However, the amplifier gain and the constant of proportionality that describes the angle-voltage characteristic of the potentiometer are never known exactly, and therefore, the actual gain in the feedback system may differ from the nominal value assumed in designing the system. Furthermore, the damping characteristics of the motor cannot be determined with absolute precision, and thus, the actual time constant of the motor response may differ from the approximate value in the specification of the system. For example, if the actual motor time constant is larger than the nominal value used in the design, the motor will respond more sluggishly than anticipated, thereby producing an effective time delay in the feedback system. As we have discussed in earlier chapters, and as we will again in Example 11.11, time delays have the effect of increasing the negative phase in the frequency response of a system, and this phase shift can have a destabilizing influence on the system. Because of the possible presence of gain and phase errors such as those we have just described, it is clearly desirable to set the amplifier gain so that there is some margin for error—that is, so that the actual system will remain stable even if it differs somewhat from the approximate model used in the design process.

In this section, we introduce one method for quantifying the margin of stability in a feedback system. To do this, we consider a closed-loop system, depicted in Figure 11.25, that has been designed to be stable based on nominal values for the forward- and feedback-path system functions. For our discussion here, we let $H(s)$ and $G(s)$ denote these nominal values. Also, since the basic concepts are identical for both continuous-time and discrete-time systems, we will again focus our development on the continuous-time case, and at the end of the section we illustrate the application of these ideas to a discrete-time example.

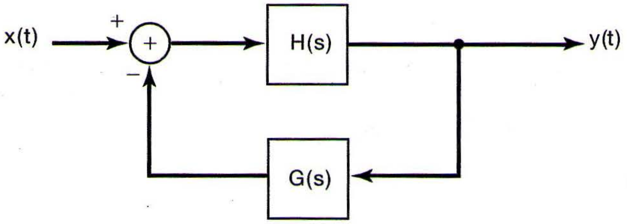


Figure 11.25 Typical feedback system designed to be stable, assuming nominal descriptions for $H(s)$ and $G(s)$.

To assess the margin of stability in our feedback system, suppose that the actual system is as depicted in Figure 11.26, where we have allowed for the possibility of a gain K and phase shift ϕ in the feedback path. In the nominal system K is unity and ϕ is zero, but in the actual system either or both may have a different value. Therefore, it is of interest to know how much variation can be tolerated in these quantities without losing closed-loop system stability. In particular, the *gain margin* of the feedback system is defined as the minimum amount of additional gain K, with $\phi = 0$, that is required so that the closed-loop system becomes unstable. Similarly, the *phase margin* is the additional amount of phase shift, with $K = 1$, that is required for the system to be unstable. By convention, the phase margin is expressed as a positive quantity; that is, it equals the magnitude of the additional negative phase shift at which the feedback system becomes unstable.

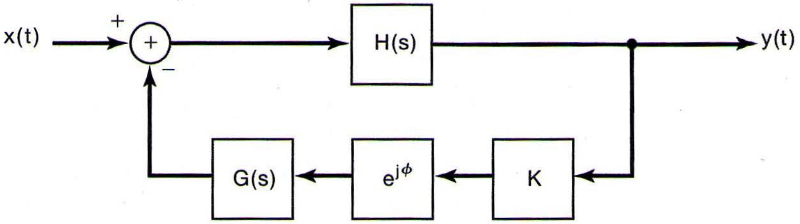


Figure 11.26 Feedback system containing possible gain and phase deviations from the nominal description depicted in Figure 11.25.

Since the closed-loop system of Figure 11.25 is stable, the system of Figure 11.26 can become unstable if, as K and ϕ are varied, at least one pole of the closed-loop system crosses the $j\omega$-axis. If a pole of the closed-loop system is on the $j\omega$-axis at, say, $\omega = \omega_0$, then at this frequency

$$1 + Ke^{-j\phi}G(j\omega_0)H(j\omega_0) = 0, \tag{11.99}$$

or

$$Ke^{-j\phi}G(j\omega_0)H(j\omega_0) = -1. \tag{11.100}$$

Note that with $K = 1$ and $\phi = 0$, by our assumption of stability for the nominal feedback system of Figure 11.25, there is no value of ω_0 for which eq. (11.100) is satisfied. The gain margin of this system is the minimum value of $K > 1$ for which eq. (11.100) has a solution for *some* ω_0 with $\phi = 0$. That is, the gain margin is the smallest value of K for which the equation

$$KG(j\omega_0)H(j\omega_0) = -1 \tag{11.101}$$

has a solution ω_0. Similarly, the phase margin is the minimum value of ϕ for which eq. (11.100) has a solution for some ω_0 when $K = 1$. In other words, the phase margin is the smallest value of $\phi > 0$ for which the equation

$$e^{-j\phi}G(j\omega_0)H(j\omega_0) = -1 \tag{11.102}$$

has a solution.

To illustrate the calculation and graphical interpretation of gain and phase margins, we consider the following example.

Example 11.9

Let

$$G(s)H(s) = \frac{4(1 + \frac{1}{2}s)}{s(1 + 2s)(1 + 0.05s + (0.125s)^2)}. \tag{11.103}$$

The Bode plot for this example is shown in Figure 11.27. Note that, as discussed in Problem 6.31, the factor of $1/j\omega$ in $G(j\omega)H(j\omega)$ contributes $-90°$ ($-\pi/2$ radians) of phase shift and a 20-dB-per-decade increase in $|G(j\omega)H(j\omega)|$. To determine the gain margin, we observe that, with $\phi = 0$, the only frequency at which eq. (11.101) can be satisfied is that for which $\sphericalangle G(j\omega_0)H(j\omega_0) = -\pi$. At this frequency, the gain margin in decibels can be identified by inspection of Figure 11.27. We first examine Figure 11.27(b) to determine the frequency ω_1 at which the angle curve crosses the line $-\pi$ radians. Locating the point at this same frequency in Figure 11.27(a) provides us with the value of $|G(j\omega_1)H(j\omega_1)|$. For eq. (11.101) to be satisfied for $\omega_0 = \omega_1$, K must equal $1/|G(j\omega_1)H(j\omega_1)|$. This value is the gain margin. As illustrated in Figure 11.27(a), the gain margin expressed in decibels can be identified as the amount the log-magnitude curve would have to be shifted up so that the curve intersects the 0-dB line at the frequency ω_1.

In a similar fashion, we can determine the phase margin. Note first that the only frequency at which eq. (11.102) can be satisfied is that for which $|G(j\omega_0)H(j\omega_0)| = 1$, or equivalently, $20 \log_{10} |G(j\omega_0)H(j\omega_0)| = 0$. To determine the phase margin, we first find the frequency ω_2 in Figure 11.27(a) at which the log-magnitude curve crosses the 0-dB line. Locating the point at this same frequency in Figure 11.27(b) then provides us with the value of $\sphericalangle G(j\omega_2)H(j\omega_2)$. For eq. (11.102) to be satisfied for $\omega_0 = \omega_2$, the angle of the left-hand side of this equation must be $-\pi$. The value of ϕ for which this is true is the phase margin. As illustrated in Figure 11.27(b), the phase margin can be identified as the amount the angle curve would have to be lowered so that the curve intersects the line $-\pi$ at the frequency ω_2.

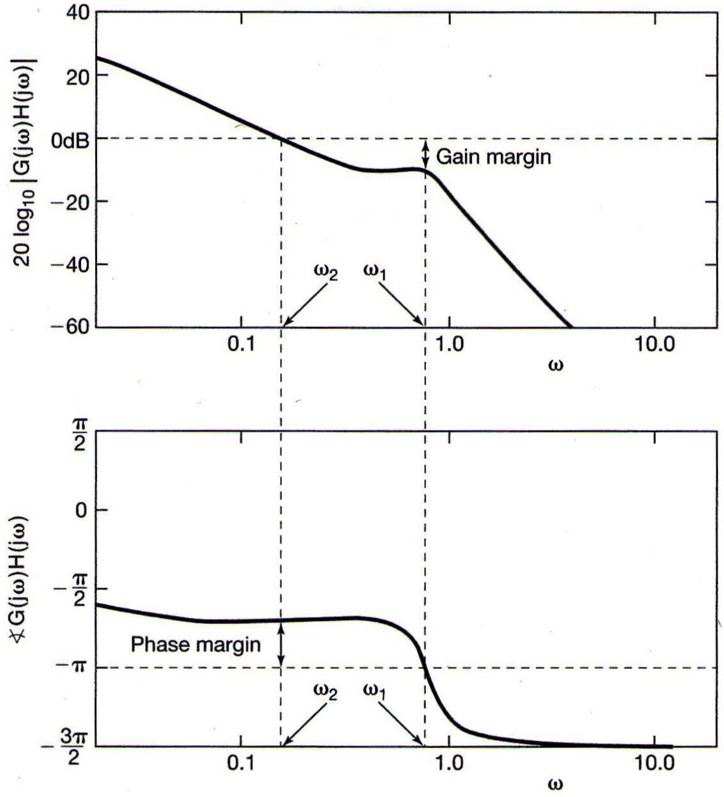


Figure 11.27 Use of Bode plots to calculate gain and phase margins for the system of Example 11.9.

In determining gain and phase margins, it is not always of interest to identify explicitly the *frequency* at which the poles will cross the $j\omega$-axis. As an alternative, we can identify the gain and phase margins from a *log magnitude-phase diagram.* For example, the log magnitude-phase diagram for the system of Figure 11.27 is shown in Figure 11.28. In this figure, we plot $20\log_{10}|G(j\omega)H(j\omega)|$ versus $\sphericalangle G(j\omega)H(j\omega)$ as ω varies from 0 to $+\infty$. Therefore, because of the conjugate symmetry of $G(j\omega)H(j\omega)$, the plot contains the same information as the Nyquist plot, in which $\mathcal{R}e\{G(j\omega)H(j\omega)\}$ is plotted versus $\mathcal{I}m\{G(j\omega)H(j\omega)\}$ for $-\infty < \omega < \infty$. As we have indicated, the phase margin can be read off by locating the intersection of the log magnitude-phase plot with the 0-dB line. That is, the phase margin is the amount of additional negative phase shift required to shift the log magnitude-phase curve so that it intersects the 0-dB line with exactly 180° (π radians) of phase shift. Similarly, the gain margin is directly obtained from the intersection of the log magnitude-phase curve with the line $-\pi$ radians, and this represents the amount of additional gain needed so that the curve crosses the line $-\pi$ with a magnitude of 0 dB.

The following examples provide several other elementary illustrations of log magnitude-phase diagrams:

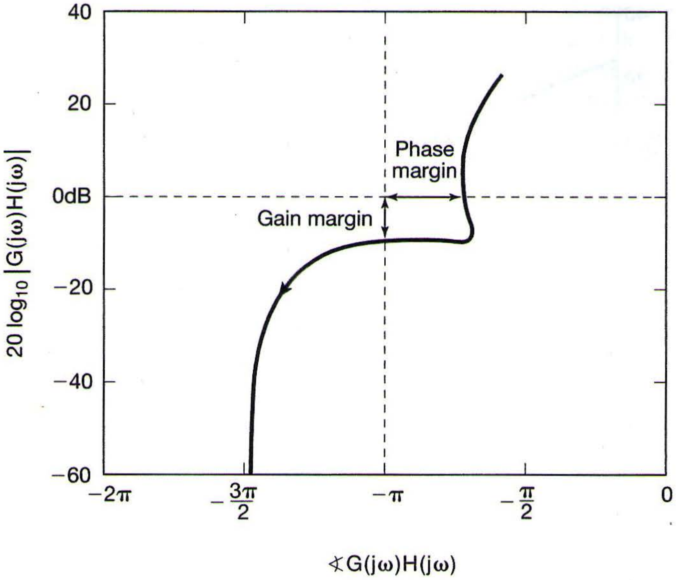


Figure 11.28 Log magnitude-phase plot for the system of Example 11.9.

Example 11.10

Let

$$G(s)H(s) = \frac{1}{\tau s + 1}, \qquad \tau > 0. \tag{11.104}$$

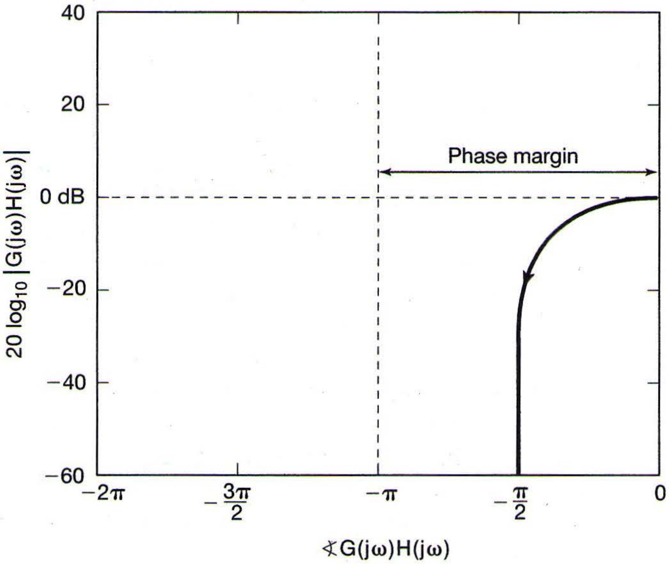


Figure 11.29 Log magnitude-phase plot for the first-order system of Example 11.10.

In this case, we obtain the log magnitude-phase plot depicted in Figure 11.29. This has a phase margin of π, and since the curve does not intersect the line $-\pi$, the system has infinite gain margin (i.e., we can increase the gain as much as we like and maintain stability). This is consistent with the conclusion that we can draw by examining the system illustrated in Figure 11.30(a). In Figure 11.30(b), we have depicted the root locus for this system with $\phi = 0$ and $K > 0$. From the figure, it is evident that the system is stable for any positive value of K. In addition, if $K = 1$ and $\phi = \pi$, so that $e^{j\phi} = -1$, the closed-loop system function for the system of Figure 11.30(a) is $1/\tau s$, which has a pole at $s = 0$, so that the system is unstable.

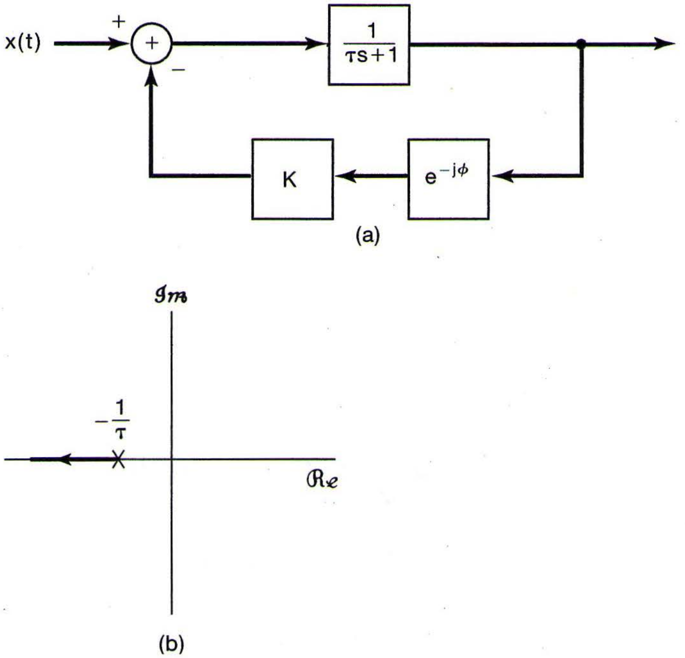


Figure 11.30 (a) First-order feedback system with possible gain and phase variations in the feedback path; (b) root locus for this system with $\phi = 0$, $K > 0$.

Example 11.11

Suppose we now consider the second-order system

$$H(s) = \frac{1}{s^2 + s + 1}, \qquad G(s) = 1. \tag{11.105}$$

The system $H(s)$ has an undamped natural frequency of 1 and a damping ratio of 0.5. The log magnitude-phase plot for this system is illustrated in Figure 11.31. Again we have infinite gain margin, but a phase margin of only $\pi/2$, since it can be shown by a straightforward calculation that $|H(j\omega)| = 1$ for $\omega = 1$, and at this frequency $\sphericalangle H(j\omega) = -\pi/2$.

We can now illustrate the type of problem that can be solved using the concepts of gain and phase margins. Suppose that the feedback system specified by eq. (11.105) cannot be realized. Rather, some unavoidable time delay is introduced into the feedback path. That is,

$$G(s) = e^{-s\tau}, \tag{11.106}$$

where τ is the time delay. What we would like to know is how small this delay must be to ensure the stability of the closed-loop system.

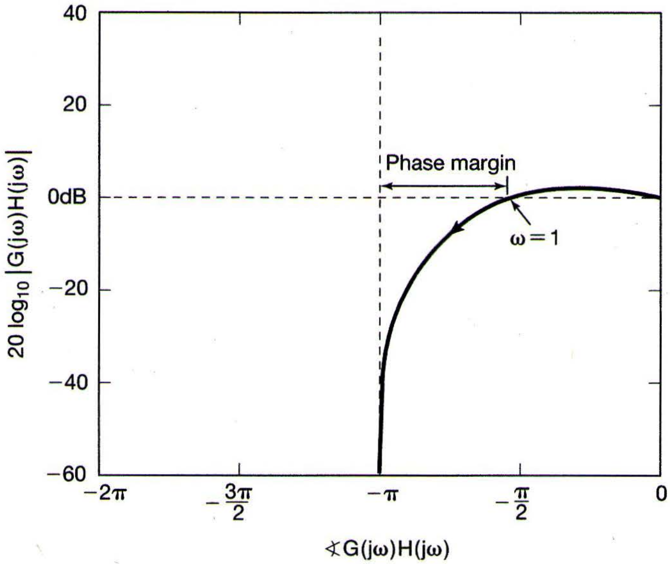


Figure 11.31 Log magnitude-phase plot for the second-order system of Example 11.11.

The first point to note is that

$$|e^{-j\omega\tau}| = 1, \tag{11.107}$$

so the delay does not change the magnitude of $H(j\omega)G(j\omega)$. On the other hand,

$$\sphericalangle e^{-j\omega\tau} = -\omega\tau \text{ radians.} \tag{11.108}$$

Thus, every point on the curve in Figure 11.31 is shifted to the *left.* The amount of the shift is proportional to the value of ω for each point on the log magnitude-phase curve.

From this discussion, we see that instability will occur once the phase margin is reduced to zero, and this will happen when the phase shift introduced by the delay is equal to $-\pi/2$ at $\omega = 1$. That is, the critical value τ^* of the time delay satisfies

$$\sphericalangle e^{-j\tau^*} = -\tau^* = -\frac{\pi}{2}, \tag{11.109}$$

or (assuming that the units of ω are radians/second)

$$\tau^* \approx 1.57 \text{ seconds.} \tag{11.110}$$

Thus, for any time delay $\tau < \tau^*$, the system remains stable.

Example 11.12

Consider again the acoustic feedback system discussed in Section 11.2.6 and Example 11.7. Here, we assume that the system of Figure 11.8 has been designed with $K_1K_2 < 1$, so that the closed-loop system is stable. In this case, the log magnitude-phase plot for

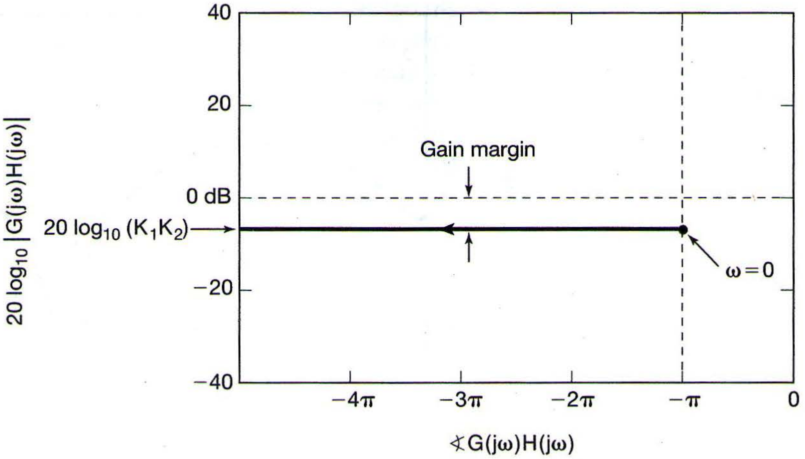


Figure 11.32 Log magnitude-phase plot for Example 11.12.

$G(s)H(s) = K_1K_2e^{-(sT+j\pi)}$ is illustrated in Figure 11.32. From the figure, we see that the system has infinite phase margin and a gain margin in decibels of $-20 \log_{10}(K_1K_2)$ (i.e., this is precisely the gain factor that, when multiplied by K_1K_2, equals 1).

As indicated at the start of the section, the definitions of the gain and phase margin are the same for discrete-time feedback systems as for continuous-time systems. Specifically, if we have a stable discrete-time feedback system, the gain margin is the minimum amount of additional gain required in the feedback system such that the closed-loop system becomes unstable. Similarly, the phase margin is the minimum amount of additional negative phase shift required for the feedback system to be unstable. The following example illustrates the graphical calculation of phase and gain margins for a discrete-time feedback system; the procedure is essentially the same as for continuous-time systems.

Example 11.13

In this example, we illustrate the concept of gain and phase margin for the discrete-time feedback system shown in Figure 11.33. Here,

$$G(z)H(z) = \frac{\frac{7\sqrt{2}}{4}z^{-1}}{1 - \frac{7\sqrt{2}}{8}z^{-1} + \frac{49}{64}z^{-2}}, \tag{11.111}$$

and by direct calculation we can check that the feedback system is stable for $K = 1$ and $\phi = 0$. In Figure 11.34, we have displayed the log magnitude-phase diagram for

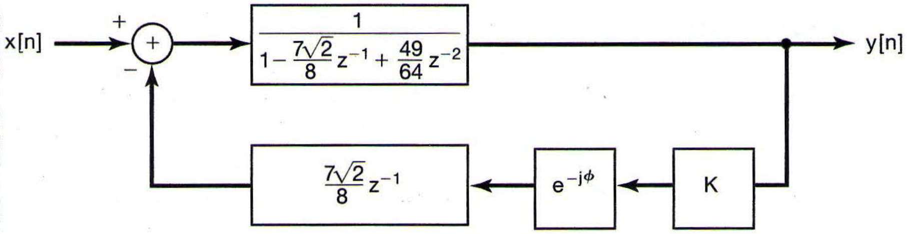

x[n]
+
−
+
$\frac{1}{1-\frac{7\sqrt{2}}{8}z^{-1}+\frac{49}{64}z^{-2}}$
y[n]
$\frac{7\sqrt{2}}{8}z^{-1}$
$e^{-j\phi}$
K

Figure 11.33 Discrete-time feedback system of Example 11.13.

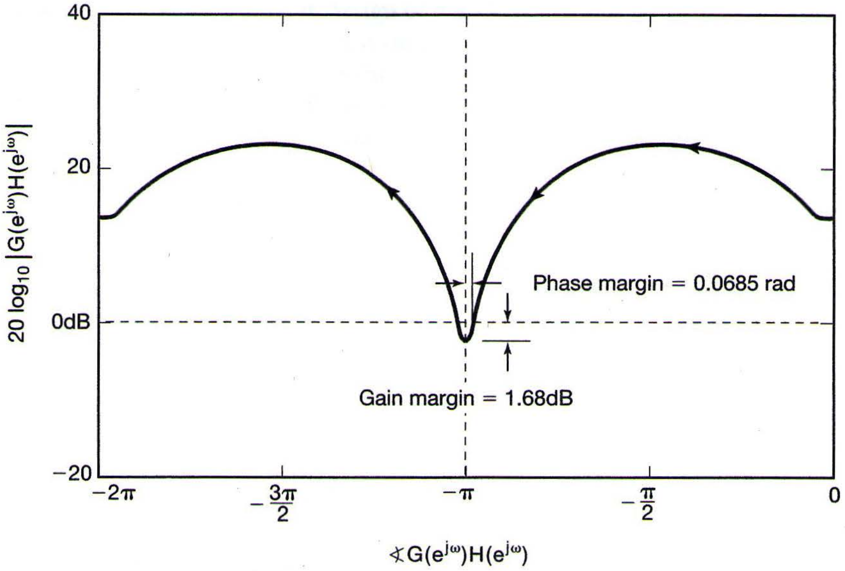


Figure 11.34 Log magnitude-phase diagram for the discrete-time feedback system of Example 11.13.

the system; that is, we have plotted $20 \log_{10} |G(e^{j\omega})H(e^{j\omega})|$ versus $\sphericalangle G(e^{j\omega})H(e^{j\omega})$ as ω varies from 0 to 2π. The system has a gain margin of 1.68 dB and a phase margin of 0.0685 radians (3.93°).

In concluding this section, it should be stressed that the gain margin is the *minimum* value of gain that moves one or more of the closed-loop poles onto the $j\omega$-axis in continuous time or the unit circle in discrete time and, consequently, causes the system to become unstable. It is important to note, however, that this does *not* imply that the system is unstable for *all* values of gain above the value specified by the gain margin. For example, as illustrated in Problem 11.47, as K increases, the root locus may move from the left-half plane into the right-half plane and then cross back into the left-half plane. The gain margin provides us with the information about how much the gain can be increased until the poles *first* reach the $j\omega$-axis, but it tells us nothing about the possibility that the system may again be stable for even larger values of the gain. To obtain such information, we must either refer to the root locus or use the Nyquist stability criterion. (See Problem 11.47.).[4]

11.6 SUMMARY

In this chapter, we have examined a number of the applications and several techniques for the analysis of feedback systems. We have seen how the use of Laplace and z-transforms allows us to analyze these systems algebraically and graphically. In Section 11.2 we indicated several of the applications of feedback, including the design of inverse systems, the stabilization of unstable systems, and the design of tracking systems. We also saw that feedback can destabilize, as well as stabilize, a system.

---

[4] For detailed discussions of this point and also of gain and phase margins and log magnitude-phase diagrams in general, see the texts on feedback listed in the bibliography at the end of the book.

In Section 11.3, we described the root-locus method for plotting the poles of the closed-loop system as a function of a gain parameter. Here, we found that the geometric evaluation of the phase of a rational Laplace transform or z-transform allowed us to gain a significant amount of insight into the properties of the root locus. These properties often permit us to obtain a reasonably accurate sketch of the root locus without performing complex calculations.

In contrast to the root-locus method, the Nyquist criterion of Section 11.4 is a technique for determining the stability of a feedback system, again as a function of a variable gain, *without* obtaining a detailed description of the location of the closed-loop poles. The Nyquist criterion is applicable to nonrational system functions and thus can be used when all that is available are experimentally determined frequency responses. The same is true of the gain and phase margins described in Section 11.5. These quantities provide a measure of the margin of stability in a feedback system and therefore are of importance to designers in that they allow them to determine how robust a feedback system is to discrepancies between estimates of the forward- and feedback-path system functions and their actual values.

Chapter 11 Problems

The first section of problems belongs to the basic category, and the answers are provided in the back of the book. The remaining three sections contain problems belonging to the basic, advanced, and extension categories, respectively.

BASIC PROBLEMS WITH ANSWERS

11.1. Consider the interconnection of discrete-time LTI systems shown in Figure P11.1. Express the overall system function for this interconnection in terms of $H_0(z)$, $H_1(z)$, and $G(z)$.

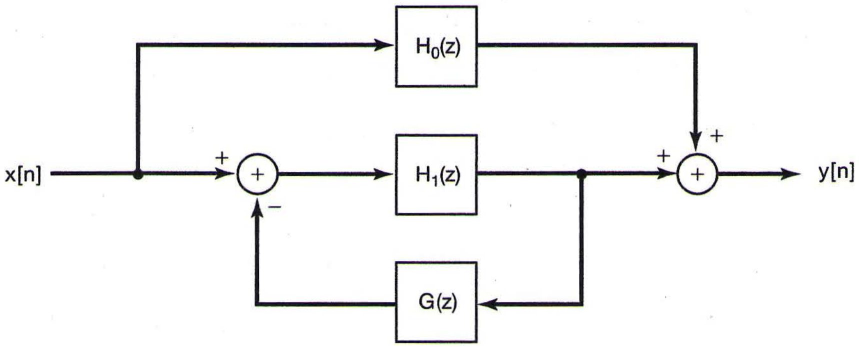


Figure P11.1

11.2. Consider the interconnection of continuous-time LTI systems shown in Figure P11.2. Express the overall system function for this interconnection in terms of $H_1(s)$, $H_2(s)$, $G_1(s)$, and $G_2(s)$.

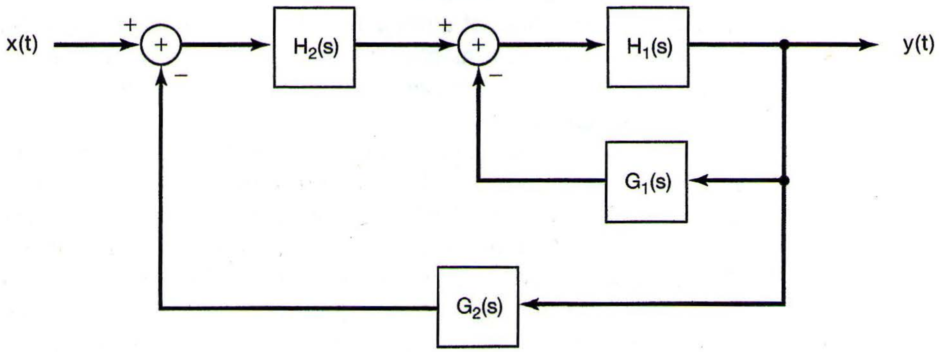


Figure P11.2

11.3. Consider the continuous-time feedback system depicted in Figure 11.3(a) with

$$H(s) = \frac{1}{s-1} \quad \text{and} \quad G(s) = s - b.$$

For what real values of b is the feedback system stable?

11.4. A causal LTI system S with input $x(t)$ and output $y(t)$ is represented by the differential equation

$$\frac{d^2y(t)}{dt^2} + \frac{dy(t)}{dt} + y(t) = \frac{dx(t)}{dt}.$$

S is to be implemented using the feedback configuration of Figure 11.3(a) with $H(s) = 1/(s+1)$. Determine $G(s)$.

11.5. Consider the discrete-time feedback system depicted in Figure 11.3(b) with

$$H(z) = \frac{1}{1 - \frac{1}{2}z^{-1}} \quad \text{and} \quad G(z) = 1 - bz^{-1}.$$

For what real values of b is the feedback system stable?

11.6. Consider the discrete-time feedback system depicted in Figure 11.3(b) with

$$H(z) = 1 - z^{-N} \quad \text{and} \quad G(z) = \frac{z^{-1}}{1 - z^{-N}}.$$

Is this system IIR or FIR?

11.7. Suppose the closed-loop poles of a feedback system satisfy

$$\frac{1}{(s+2)(s+3)} = -\frac{1}{K}.$$

Use the root-locus method to determine the values of K for which the feedback system is guaranteed to be stable.

11.8. Suppose the closed-loop poles of a feedback system satisfy

$$\frac{s-1}{(s+1)(s+2)} = -\frac{1}{K}.$$

Use the root-locus method to determine the negative values of K for which the feedback system is guaranteed to be stable.

11.9. Suppose the closed-loop poles of a feedback system satisfy

$$\frac{(s+1)(s+3)}{(s+2)(s+4)} = -\frac{1}{K}.$$

Use the root-locus method to determine whether there are any values of the adjustable gain K for which the system's impulse response has an oscillatory component of the form $e^{-at}\cos(\omega_0 t + \phi)$, where $\omega_0 \neq 0$.

11.10. The root locus corresponding to $G(s)H(s) = -1/K$ is illustrated in Figure P11.10. In this figure, the start ($K = 0$) and end of each branch of the root locus are marked by a '●' symbol. Specify the poles and zeros of $G(s)H(s)$.

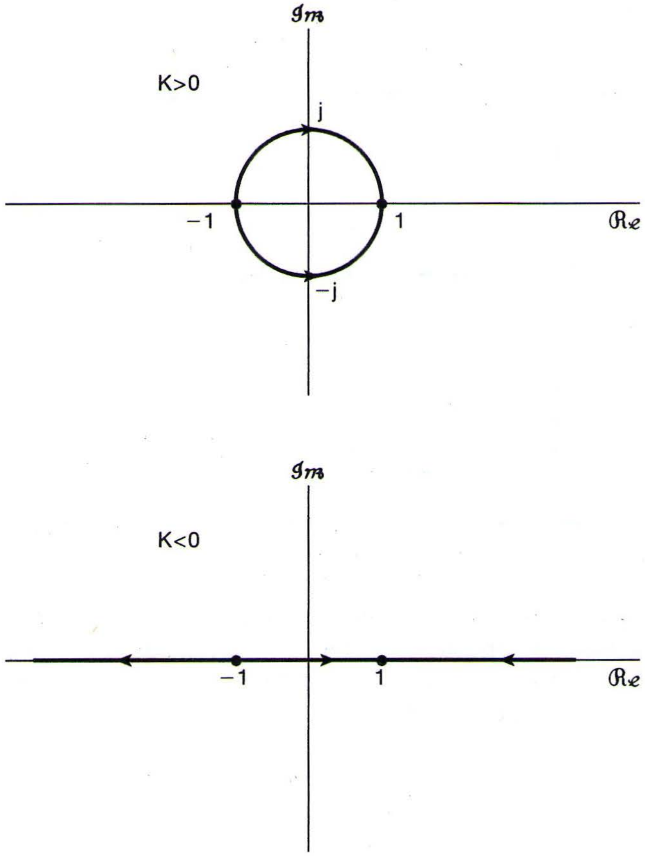


Figure P11.10

11.11. Suppose the closed-loop poles of a discrete-time feedback system satisfy

$$\frac{z^{-2}}{(1-\frac{1}{2}z^{-1})(1+\frac{1}{2}z^{-1})} = -\frac{1}{K}.$$

Using the root-locus method, determine the positive values of K for which this system is stable.

11.12. Each of the four locations $z = 1/2$, $z = 1/4$, $z = 0$, and $z = -1/2$ is a single-order pole or zero of $G(z)H(z)$. Furthermore, $G(z)H(z)$ is known to have only two poles. What information can you deduce about the poles and zeros of $G(z)H(z)$ from the fact that for all K, the root locus corresponding to

$$G(z)H(z) = -\frac{1}{K}$$

is on the real axis.

11.13. Consider the block diagram of Figure P11.13 for a discrete-time system. Use the root-locus method to determine the values of K for which the system is guaranteed to be stable.

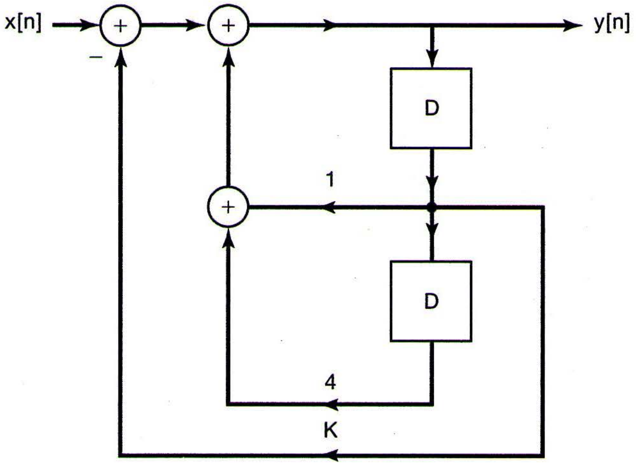


Figure P11.13

11.14. Let C be a closed path that lies on the unit circle in the p-plane and that is traversed in the clockwise direction in order to evaluate $W(p)$. For each of the following expressions for $W(p)$, determine the net number of times the plot of $W(p)$ encircles the origin in a clockwise direction:

(a) $W(p) = \dfrac{(1-\frac{1}{2}p^{-1})}{(1-\frac{1}{4}p^{-1})}$

(b) $W(p) = \dfrac{(1-2p^{-1})}{(1-\frac{1}{2}p^{-1})(1-2p^{-1}+4p^{-2})}$

11.15. Consider a continuous-time feedback system whose closed-loop poles satisfy

$$G(s)H(s) = \frac{1}{(s+1)} = -\frac{1}{K}.$$

Use the Nyquist plot and the Nyquist stability criterion to determine the range of values of K for which the closed-loop system is stable. *Hint:* In sketching the Nyquist plot, you may find it useful to sketch the corresponding Bode plot first. It also is helpful to determine the values of ω for which $G(j\omega)H(j\omega)$ is real.

11.16. Consider a continuous-time feedback system whose closed-loop poles satisfy

$$G(s)H(s) = \frac{1}{(s+1)(s/10+1)} = -\frac{1}{K}.$$

Use the Nyquist plot and the Nyquist stability criterion to determine the range of values of K for which the closed-loop system is stable.

11.17. Consider a continuous-time feedback system whose closed-loop poles satisfy

$$G(s)H(s) = \frac{1}{(s+1)^4} = -\frac{1}{K}.$$

Use the Nyquist plot and the Nyquist stability criterion to determine the range of values of K for which the closed-loop system is stable.

11.18. Consider a discrete-time feedback system whose closed-loop poles satisfy

$$G(z)H(z) = z^{-3} = -\frac{1}{K}.$$

Use the Nyquist plot and the Nyquist stability criterion to determine the range of values of K for which the closed-loop system is stable.

11.19. Consider a feedback system, either in continuous-time or discrete-time, and suppose that the Nyquist plot for the system passes through the point $-1/K$. Is the feedback system stable or unstable for this value of the gain? Explain your answer.

11.20. Consider the basic continuous-time feedback system of Figure 11.3(a). Determine the phase and gain margin for the following specification of $H(s)$ and $G(s)$:

$$H(s) = \frac{s+1}{s^2+s+1}, \qquad G(s) = 1.$$

BASIC PROBLEMS

11.21. Consider the feedback system of Figure P11.21. Find the closed-loop poles and zeros of this system for the following values of K:

(i) $K = 0.1$
(ii) $K = 1$
(iii) $K = 10$
(iv) $K = 100$

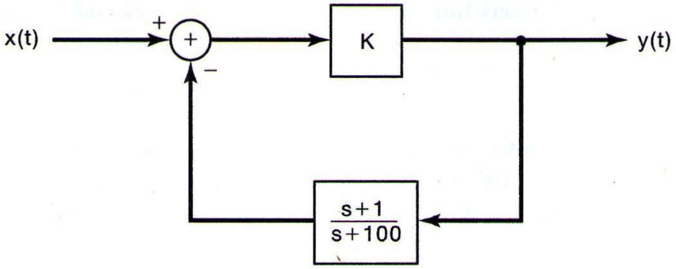


Figure P11.21

11.22. Consider the basic feedback system of Figure 11.3(a). Determine the closed-loop system impulse response for each of the following specifications of the system functions in the forward and feedback paths:

(a) $H(s) = \frac{1}{(s+1)(s+3)}$, $G(s) = 1$

(b) $H(s) = \frac{1}{s+3}$, $G(s) = \frac{1}{s+1}$

(c) $H(s) = \frac{1}{2}$, $G(s) = e^{-s/3}$

11.23. Consider the basic feedback systems of Figure 11.3(b). Determine the closed-loop system impulse response for each of the following specifications of the system functions in the forward and feedback paths:

(a) $H(z) = \frac{z^{-1}}{1-\frac{1}{2}z^{-1}}$, $G(z) = \frac{2}{3} - \frac{1}{6}z^{-1}$

(b) $H(z) = \frac{2}{3} - \frac{1}{6}z^{-1}$, $G(z) = \frac{z^{-1}}{1-\frac{1}{2}z^{-1}}$

11.24. Sketch the root loci for $K > 0$ and $K < 0$ for each of the following:

(a) $G(s)H(s) = \frac{1}{s+1}$

(b) $G(s)H(s) = \frac{1}{(s-1)(s+3)}$

(c) $G(s)H(s) = \frac{1}{s^2+s+1}$

(d) $G(s)H(s) = \frac{s+1}{s^2}$

(e) $G(s)H(s) = \frac{(s+1)^2}{s^3}$

(f) $G(s)H(s) = \frac{s^2+2s+2}{s^2(s-1)}$

(g) $G(s)H(s) = \frac{(s+1)(s-1)}{s(s^2+2s+2)}$

(h) $G(s)H(s) = \frac{(1-s)}{(s+2)(s+3)}$

11.25. Sketch the root loci for $K > 0$ and $K < 0$ for each of the following:

(a) $G(z)H(z) = \frac{z-1}{z^2-\frac{1}{4}}$

(b) $G(z)H(z) = \frac{2}{z^2-\frac{1}{4}}$

(c) $G(z)H(z) = \frac{z^{-1}(1+z^{-1})}{1-\frac{1}{4}z^{-2}}$

(d) $G(z)H(z) = z^{-1} - z^{-2}$

(e) $G(z)H(z)$ is the system function of the causal LTI system described by the difference equation

$$y[n] - 2y[n-1] = x[n-1] - x[n-2].$$

11.26. Consider a feedback system with

$$G(s)H(s) = \frac{(s-a)(s-b)}{s(s+3)(s+6)}.$$

Sketch the root locus for $K > 0$ and $K < 0$ for the following values of a and b:

(a) $a = 1, \quad b = 2$ **(b)** $a = -2, \quad b = 2$ **(c)** $a = -4, \quad b = 2$
(d) $a = -7, \quad b = 2$ **(e)** $a = -1, \quad b = -2$ **(f)** $a = -4, \quad b = -2$
(g) $a = -7, \quad b = -2$ **(h)** $a = -5, \quad b = -4$ **(i)** $a = -7, \quad b = -4$
(j) $a = -7, \quad b = -8$

11.27. Consider a feedback system with

$$H(s) = \frac{s+2}{s^2+2s+4}, \qquad G(s) = K.$$

(a) Sketch the root locus for $K > 0$.
(b) Sketch the root locus for $K < 0$.
(c) Find the smallest positive value of K for which the closed-loop impulse response does not exhibit any oscillatory behavior.

11.28. Sketch the Nyquist plot for each of the following specifications of $G(s)H(s)$, and use the continuous-time Nyquist criterion to determine the range of values of K (if any such range exists) for which the closed-loop system is stable. *Note:* In sketching the Nyquist plots, you may find it useful to sketch the corresponding Bode plots first. It also is helpful to determine the values of ω for which $G(j\omega)H(j\omega)$ is real.

(a) $G(s)H(s) = \frac{1}{s-1}$ **(b)** $G(s)H(s) = \frac{1}{s^2-1}$
(c) $G(s)H(s) = \frac{1}{(s+1)^2}$ **(d)** $G(s)H(s) = \frac{1}{(s+1)^3}$
(e) $G(s)H(s) = \frac{1-s}{(s+1)^2}$ **(f)** $G(s)H(s) = \frac{s+1}{(s-1)^2}$
(g) $G(s)H(s) = \frac{s+1}{s^2-4}$ **(h)** $G(s)H(s) = \frac{1}{s^2+2s+2}$
(i) $G(s)H(s) = \frac{s+1}{s^2-2s+2}$ **(j)** $G(s)H(s) = \frac{s+1}{(s+100)(s-1)^2}$
(k) $G(s)H(s) = \frac{s^2}{(s+1)^3}$

11.29. Consider the basic continuous-time feedback system of Figure 11.3(a). Sketch the log magnitude-phase diagram, and roughly determine the phase and gain margin, for each of the following choices of $G(s)$ and $H(s)$. You may find it useful to use the straight-line approximations to the Bode plots developed in Chapter 6 to aid you in sketching the log magnitude-phase diagrams. Be careful, however, to take into account how the actual frequency response deviates from its approximation near break frequencies when there are underdamped second-order terms present. (See Section 6.5.2.)

(a) $H(s) = \frac{10s+1}{s^2+s+1}$, $G(s) = 1$
(b) $H(s) = \frac{s/10+1}{s^2+s+1}$, $G(s) = 1$
(c) $H(s) = \frac{1}{(s+1)^2(s+10)}$, $G(s) = 100$
(d) $H(s) = \frac{1}{(s+1)^3}$, $G(s) = \frac{1}{s+1}$
(e) $H(s) = \frac{1-s}{(s+1)(s+10)}$, $G(s) = 1$

(f) $H(s) = \frac{1-s/100}{(s+1)^2}$, $G(s) = \frac{10s+1}{s/10+1}$

(g) $H(s) = \frac{1}{s(s+1)}$, $G(s) = \frac{1}{s+1}$

Note: Your sketch for part (g) should reflect the fact that for this feedback system $|G(j\omega)H(j\omega)| \to \infty$ as $\omega \to 0$; what is the phase of $G(j\omega)H(j\omega)$ for $\omega = 0^+$, i.e., for ω an infinitesimal amount larger than 0?

11.30. Sketch the Nyquist plot for each of the following specifications of $G(z)H(z)$, and use the discrete-time Nyquist criterion to determine the range of values of K (if any such range exists) for which the closed-loop system is stable. [*Note:* In sketching the Nyquist plots, you may find it useful to first sketch the magnitude and phase plots as a function of frequency or at least calculate $|G(e^{j\omega})H(e^{j\omega})|$ and $\sphericalangle G(e^{j\omega})H(e^{j\omega})$ at several points. Also, it is helpful to determine the values of ω for which $G(e^{j\omega})H(e^{j\omega})$ is real.]

(a) $G(z)H(z) = \frac{1}{z-\frac{1}{2}}$ **(b)** $G(z)H(z) = \frac{1}{z-2}$

(c) $G(z)H(z) = z^{-1}$ **(d)** $G(z)H(z) = z^{-2}$

(e) $G(z)H(z) = \frac{1}{(z+\frac{1}{2})(z-\frac{3}{2})}$ **(f)** $G(z)H(z) = \frac{z-\sqrt{3}}{z(z+1/\sqrt{3})}$

(g) $G(z)H(z) = \frac{1}{z^2-z+\frac{1}{3}}$ **(h)** $G(z)H(z) = \frac{z-\frac{1}{2}}{z(z-2)}$

(i) $G(z)H(z) = \frac{(z+1)^2}{z^3}$

11.31. Consider the basic discrete-time system in Figure 11.3(b). Sketch the log magnitude-phase diagram, and roughly determine the phase and gain margin, for each of the following choices of $G(z)$ and $H(z)$. You may find it useful to determine the values of ω for which either $|G(e^{j\omega})H(e^{j\omega})| = 1$ or $\sphericalangle G(e^{j\omega}) = -\pi$.

(a) $H(z) = z^{-1}$, $G(z) = \frac{1}{2}$

(b) $H(z) = \frac{z^{-1}}{1-\frac{1}{2}z^{-1}}$, $G(z) = \frac{1}{2}$

(c) $H(z) = \frac{1}{(1-\frac{1}{2}z^{-1})(1+\frac{1}{2}z^{-1})}$, $G(z) = z^{-2}$

(d) $H(z) = \frac{2}{z-2}$, $G(z) = 1$

(e) $H(z) = \frac{1}{z+\frac{1}{2}}$, $G(z) = \frac{1}{z-\frac{3}{2}}$

(f) $H(z) = \frac{1}{z+\frac{1}{2}}$, $G(z) = 1 - \frac{3}{2}z^{-1}$

(g) $H(z) = \frac{\frac{1}{2}}{z^2-z+\frac{1}{3}}$, $G(z) = 1$

(h) $H(z) = \frac{1}{z-1}$, $G(z) = \frac{1}{4}z^{-1}$

(*Note:* Your sketch for part (h) should reflect the fact that, for this feedback system, $G(z)H(z)$ has a pole at $z = 1$; what are the values of $\sphericalangle G(e^{j\omega})H(e^{j\omega})$ for $e^{j\omega}$ just on either side of the point $z = 1$?)

ADVANCED PROBLEMS

11.32. **(a)** Consider the feedback system of Figure 11.10(b) with

$$H(s) = \frac{N_1(s)}{D_1(s)}, \quad G(s) = \frac{N_2(s)}{D_2(s)}. \tag{P11.32–1}$$

Assume that there is no pole-zero cancellation in the product $G(s)H(s)$. Show that the zeros of the closed-loop system function consist of the zeros of $H(s)$ and the poles of $G(s)$.

(b) Use the result of part (a) together with the appropriate property of the root locus to confirm that, with $K = 0$, the closed-loop system zeros are the zeros of $H(s)$ and the closed-loop poles are the poles of $H(s)$.

(c) While it is usual for $H(s)$ and $G(s)$ in eq. (P11.32–1) to be in reduced form [i.e., the polynomials $N_1(s)$ and $D_1(s)$ have no common factors, and the same is true of $N_2(s)$ and $D_2(s)$], it may happen that $N_1(s)$ and $D_2(s)$ have common factors or $N_2(s)$ and $D_1(s)$ have common factors. To see what occurs when such common factors are present, let $p(s)$ denote the greatest common factor of $N_1(s)$ and $D_2(s)$. That is,

$$\frac{N_1(s)}{p(s)} \quad \text{and} \quad \frac{D_2(s)}{p(s)}$$

are both polynomials and have *no* common factors. Similarly,

$$\frac{N_2(s)}{q(s)} \quad \text{and} \quad \frac{D_1(s)}{q(s)}$$

are polynomials and have no common factors. Show that the closed-loop system function can be written as

$$Q(s) = \frac{p(s)}{q(s)}\left[\frac{\hat{H}(s)}{1 + K\hat{G}(s)\hat{H}(s)}\right], \tag{P11.32–2}$$

where

$$\hat{H}(s) = \frac{N_1(s)/p(s)}{D_1(s)/q(s)}$$

and

$$\hat{G}(s) = \frac{N_2(s)/q(s)}{D_2(s)/p(s)}.$$

Therefore, from eq. (P11.32–2) and part (a), we see that the zeros of $Q(s)$ are the zeros of $p(s)$, the zeros of $\hat{H}(s)$, and the poles of $\hat{G}(s)$, while the poles of $Q(s)$ are the zeros of $q(s)$ and the solutions of

$$1 + K\hat{G}(s)\hat{H}(s) = 0. \tag{P11.32–3}$$

By construction, there is no pole-zero cancellation in the product $\hat{G}(s)\hat{H}(s)$, and thus, we can apply the root-locus method described in Section 11.3 to sketch the locations of the solutions of eq. (P11.32–3) as K is varied.

(d) Use the procedure outlined in part (c) to determine the closed-loop zeros, any closed-loop poles whose locations are independent of K, and the locus of the remaining closed-loop poles for $K > 0$ when

$$H(s) = \frac{s+1}{(s+4)(s+2)}, \qquad G(s) = \frac{s+2}{s+1}.$$

(e) Repeat part (d) for

$$H(z) = \frac{1+z^{-1}}{1-\frac{1}{2}z^{-1}}, \quad G(z) = \frac{z^{-1}}{1+z^{-1}}.$$

(f) Let

$$H(z) = \frac{z^2}{(z-2)(z+2)}, \quad G(z) = \frac{1}{z^2}.$$

(i) Sketch the root locus for $K > 0$ and for $K < 0$.
(ii) Find all the values of K for which the overall system is stable.
(iii) Find the impulse response of the closed-loop system when $K = 4$.

11.33. Consider the feedback system of Figure 11.10(a), and suppose that

$$G(s)H(s) = \frac{\prod_{k=1}^{m}(s-\beta_k)}{\prod_{k=1}^{n}(s-\alpha_k)},$$

where $m > n$.[5] In this case $G(s)H(s)$ has $m-n$ poles at infinity (see Chapter 9), and we can adapt the root-locus rules given in the text by noting that (1) there are m branches of the root locus and (2) for $K = 0$, all branches of the root locus begin at poles of $G(s)H(s)$, $m-n$ of which are at infinity. Furthermore, as $|K| \to \infty$, these branches converge to the m zeros of $G(s)H(s)$, namely, $\beta_1, \beta_2, \ldots, \beta_m$. Use these facts to assist you in sketching the root locus (for $K > 0$ and for $K < 0$) for each of the following:

(a) $G(s)H(s) = s - 1$
(b) $G(s)H(s) = (s+1)(s+2)$
(c) $G(s)H(s) = \frac{(s+1)(s+2)}{s-1}$

11.34. In Section 11.3, we derived a number of properties that can be of value in determining the root locus for a feedback system. In this problem, we develop several additional properties. We derive these properties in terms of continuous-time systems, but, as with all root-locus properties, they hold as well for discrete-time root loci. For our discussion of these properties, we refer to the basic equation satisfied by the closed-loop poles, namely,

$$G(s)H(s) = -\frac{1}{K}, \tag{P11.34–1}$$

---

[5] Note that for a continuous-time system, the condition $m > n$ implies that the system with system function $G(s)H(s)$ involves differentiation of the input. [In fact, the inverse transform of $G(s)H(s)$ includes singularity functions up to the order $m-n$.] In discrete time, if $G(z)H(z)$, written as a ratio of polynomials in z, has $m > n$, it is necessarily the system function of a noncausal system. [In fact, the inverse transform of $G(z)H(z)$ has a nonzero value at time $n-m < 0$.] Thus, the case considered in this problem is actually of interest only for continuous-time systems.

where

$$G(s)H(s) = \frac{\prod_{k=1}^{m}(s-\beta_k)}{\prod_{k=1}^{n}(s-\alpha_k)} = \frac{\sum_{k=0}^{m} b_k s^k}{\sum_{k=0}^{n} a_k s^k}. \qquad \text{(P11.34–2)}$$

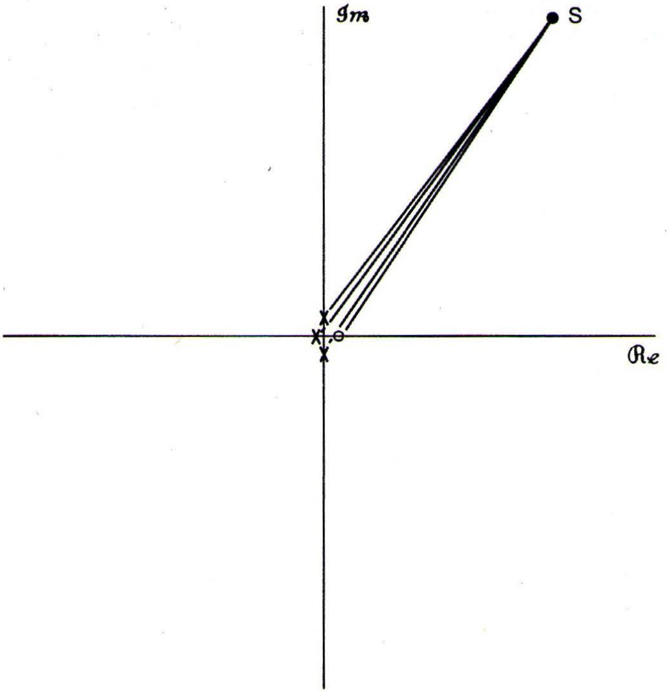


Figure P11.34

Throughout this problem, we assume that $m \le n$.

(a) From Property 2, we know that $n - m$ branches of the root locus go to zeros of $G(s)H(s)$ located at infinity. In this first part, we demonstrate that it is straightforward to determine the angles at which these branches approach infinity. Specifically, consider searching the remote part of the s-plane [i.e., the region where $|s|$ is extremely large and far from any of the poles and zeros of $G(s)H(s)$]. This region is illustrated in Figure P11.34. Use the geometry of the picture, together with the angle criterion for $K > 0$ and for $K < 0$, to deduce that:

- For $K > 0$, the $n - m$ branches of the root locus that approach infinity do so at the angles

$$\frac{(2k+1)\pi}{n-m}, \qquad k = 0, 1, \ldots, n-m-1.$$

- For $K < 0$, the $n - m$ branches of the root locus that approach infinity do so at the angles

$$\frac{2k\pi}{n-m}, \qquad k = 0, 1, \ldots, n-m-1.$$

Thus, the branches of the root locus that approach infinity do so at specified angles that are arranged symmetrically. For example, for $n - m = 3$ and $K > 0$, we see that the asymptotic angles are $\pi/3$, π, and $5\pi/3$. The result of part (a), together with one additional fact, allows us to draw in the asymptotes for the branches of the root locus that approach infinity. Specifically, all of the $n - m$ asymptotes intersect at a single point on the real axis. This is derived in the next part of the problem.

(b) (i) As a first step, consider a general polynomial equation

$$s^r + f_{r-1}s^{r-1} + \cdots + f_0 = (s - \xi_1)(s - \xi_2) \cdots (s - \xi_r) = 0.$$

Show that

$$f_{r-1} = -\sum_{i=1}^{r} \xi_i.$$

(ii) Perform long division on $1/G(s)H(s)$ to write

$$\frac{1}{G(s)H(s)} = s^{n-m} + \gamma_{n-m-1}s^{n-m-1} + \ldots. \qquad \text{(P11.34–3)}$$

Show that

$$\gamma_{n-m-1} = a_{n-1} - b_{m-1} = \sum_{k=1}^{m} \beta_k - \sum_{k=1}^{n} \alpha_k.$$

[See eq. (P11.34–2).]

(iii) Argue that the solution of eq. (P11.34–1) for *large s* is an approximate solution of the equation

$$s^{n-m} + \gamma_{n-m-1}s^{n-m-1} + \gamma_{n-m-2}s^{n-m-2} + \cdots + \gamma_0 + K = 0.$$

(iv) Use the results of (i)–(iii) to deduce that the sum of the $n - m$ closed-loop poles that approach infinity is asymptotically equal to

$$b_{m-1} - a_{n-1}.$$

Thus, the center of gravity of these $n - m$ poles is

$$\frac{b_{m-1} - a_{n-1}}{n-m},$$

which does not depend on K. Consequently, we have $n - m$ closed-loop poles that approach $|s| = \infty$ at evenly spaced angles and that have a center of gravity that is independent of K. From this, we can deduce that:

The asymptotes of the $n-m$ branches of the root locus that approach infinity intersect at the point

$$\frac{b_{m-1}-a_{n-1}}{n-m}=\frac{\sum_{k=1}^{n}\alpha_k-\sum_{k=1}^{m}\beta_k}{n-m}.$$

This point of intersection of the asymptotes is the same for $K>0$ and $K<0$.

(c) Suppose that

$$G(s)H(s)=\frac{1}{(s+1)(s+3)(s+5)}.$$

(i) What are the asymptotic angles for the closed-loop poles that approach infinity for $K>0$ and for $K<0$?
(ii) What is the point of intersection of the asymptotes?
(iii) Draw in the asymptotes, and use them to help you sketch the root locus for $K>0$ and for $K<0$.

(d) Repeat part (c) for each of the following:
(i) $G(s)H(s)=\frac{s+1}{s(s+2)(s+4)}$
(ii) $G(s)H(s)=\frac{1}{s^4}$
(iii) $G(s)H(s)=\frac{1}{s(s+1)(s+5)(s+6)}$
(iv) $G(s)H(s)=\frac{1}{(s+2)^2(s-1)^2}$
(v) $G(s)H(s)=\frac{s+3}{(s+1)(s^2+2s+2)}$
(vi) $G(s)H(s)=\frac{s+1}{(s+2)^2(s^2+2s+2)}$
(vii) $G(s)H(s)=\frac{s+1}{(s+100)(s-1)(s-2)}$

(e) Use the result of part (a) to explain why the following statement is true: For any continuous-time feedback system, with $G(s)H(s)$ given by eq. (P11.34–2), if $n-m\geq 3$, we can make the closed-loop system unstable by choosing $|K|$ large enough.

(f) Repeat part (c) for the discrete-time feedback system specified by

$$G(z)H(z)=\frac{z^{-3}}{(1-z^{-1})(1+\frac{1}{2}z^{-1})}.$$

(g) Explain why the following statement is true: For any discrete-time feedback system with

$$G(z)H(z)=\frac{z^m+b_{m-1}z^{m-1}+\cdots+b_0}{z^n+a_{n-1}z^{n-1}+\cdots+a_0},$$

if $n>m$, we can make the closed-loop system unstable by choosing $|K|$ large enough.

11.35. (a) Consider again the feedback system of Example 11.2:

$$G(s)H(s)=\frac{s-1}{(s+1)(s+2)}.$$

The root locus for $K < 0$ is plotted in Figure 11.14(b). For some value of K, the closed-loop poles are on the $j\omega$-axis. Determine this value of K and the corresponding locations of the closed-loop poles by examining the real and imaginary parts of the equation

$$G(j\omega)H(j\omega) = -\frac{1}{K},$$

which must be satisfied if the point $s = j\omega$ is on the root locus for any given values of K. Use this result plus the analysis in Example 11.2 to find the full range of values of K (positive and negative) for which the closed-loop system is stable.

(b) Note that the feedback system is unstable for $|K|$ sufficiently large. Explain why this is true in general for continuous-time feedback systems for which $G(s)H(s)$ has a zero in the right-half plane and for discrete-time feedback systems for which $G(z)H(z)$ has a zero outside the unit circle.

11.36. Consider a continuous-time feedback system with

$$G(s)H(s) = \frac{1}{s(s+1)(s+2)}. \tag{P11.36–1}$$

(a) Sketch the root locus for $K > 0$ and for $K < 0$. (*Hint:* The results of Problem 11.34 are useful here.)

(b) If you have sketched the locus correctly, you will see that for $K > 0$, two branches of the root locus cross the $j\omega$-axis, passing from the left-half plane into the right-half plane. Consequently, we can conclude that the closed-loop system is stable for $0 < K < K_0$, where K_0 is the value of the gain for which the two branches of the root locus intersect the $j\omega$-axis. Note that the sketch of the root locus does not by itself tell us what the value of K_0 is or the exact point on the $j\omega$-axis where the branches cross. As in Problem 11.35, determine K_0 by solving the pair of equations obtained as the real and imaginary parts of

$$G(j\omega)H(j\omega) = -\frac{1}{K_0}. \tag{P11.36–2}$$

Determine the corresponding two values of ω (which are the negatives of each other, since poles occur in complex-conjugate pairs).

From your root-locus sketches in part (a), note that there is a segment of the real axis between two poles which is on the root locus for $K > 0$, and a different segment is on the locus for $K < 0$. In both cases, the root locus breaks off from the real axis at some point. In the next part of this problem, we illustrate how one can calculate these breakaway points.

(c) Consider the equation denoting the closed-loop poles:

$$G(s)H(s) = -\frac{1}{K}. \tag{P11.36–3}$$

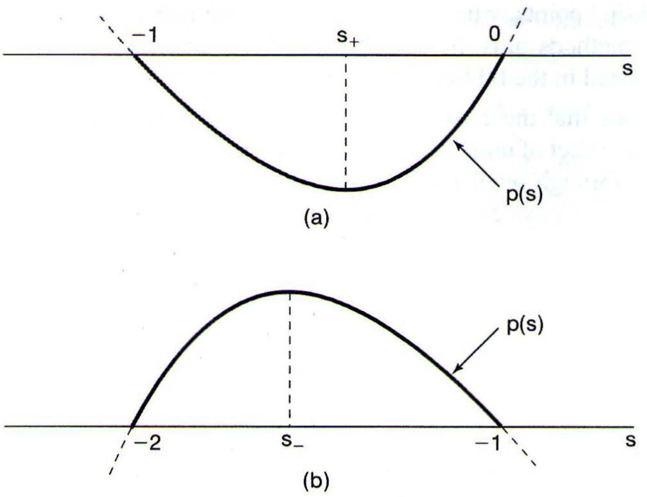


Figure P11.36

Using eq. (P11.36–1), show that an equivalent equation for the closed loop poles is

$$p(s) = s^3 + 3s^2 + 2s = -K. \tag{P11.36–4}$$

Consider the segment of the real axis between 0 and -1. This segment is on the root locus for $K \geq 0$. For $K = 0$, two branches of the locus begin at 0 and -1 and approach each other as K is increased.

(i) Use the facts stated, together with eq. (P11.36–4), to explain why the function $p(s)$ has the form shown in Figure P11.36(a) for $-1 \leq s \leq 0$ and why the point s_+ where the minimum occurs is the breakaway point (i.e., it is the point where the two branches of the $K > 0$ locus break from the segment of the real axis between -1 and 0).

Similarly, consider the root locus for $K < 0$ and, more specifically, the segment of the real axis between -1 and -2 that is part of this locus. For $K = 0$, two branches of the root locus begin at -1 and -2, and as K is decreased, these poles approach each other.

(ii) In an analogous fashion to that used in part (i), explain why the function $p(s)$ has the form shown in Figure P11.36(b) and why the point s_- where the maximum occurs is the breakaway point for $K < 0$.

Thus, the breakaway points correspond to the the maxima and minima of $p(s)$ as s ranges over the negative real line.

(iii) The points at which $p(s)$ has a maximum or minimum are the solutions of the equation

$$\frac{dp(s)}{ds} = 0.$$

Use this fact to find the breakaway points s_+ and s_-, and then use eq. (P11.36–4) to find the gains at which these points are closed-loop poles.

In addition to the method illustrated in part (c), there are other, partially analytical, partially graphical methods for determining breakaway points. It is also possible to use a procedure similar to the one just illustrated in part (c) to find the "break-in" points, where two branches of the root locus merge onto the real axis. These methods plus the one illustrated are described in advanced texts such as those listed in the bibliography at the end of the book.

11.37. One issue that must always be taken into account by the system designer is the possible effect of unmodeled aspects of the system one is attempting to stabilize or modify through feedback. In this problem, we provide an illustration of why this is the case. Consider a continuous-time feedback system, and suppose that

$$H(s) = \frac{1}{(s+10)(s-2)} \tag{P11.37–1}$$

and

$$G(s) = K. \tag{P11.37–2}$$

(a) Use root-locus techniques to show that the closed-loop system will be stable if K is chosen large enough.

(b) Suppose that the system we are trying to stabilize by feedback actually has a system function

$$H(s) = \frac{1}{(s+10)(s-2)(10^{-3}s+1)}. \tag{P11.37–3}$$

The added factor can be thought of as representing a first-order system in cascade with the system of eq. (P11.37–1). Note that the time constant of the added first order system is extremely small and thus will appear to have a step response that is almost instantaneous. For this reason, one often neglects such factors in order to obtain simpler and more tractable models that capture all of the important characteristics of the system. However, one must still keep these neglected dynamics in mind in obtaining a useful feedback design. To see why this is the case, show that if $G(s)$ is given by eq. (P11.37–2) and $H(s)$ is as in eq. (P11.37–3), then the closed-loop system will be unstable if K is chosen *too* large. *Hint:* See Problem 11.34.

(c) Use root-locus techniques to show that if

$$G(s) = K(s+100),$$

then the feedback system will be stable for all values of K sufficiently large if $H(s)$ is given by eq. (P11.37–1) *or* eq. (P11.37–3).

11.38. Consider the feedback system of Figure 11.3(b) with

$$H(z) = \frac{Kz^{-1}}{1-z^{-1}}$$

and

$$G(z) = 1 - az^{-1}.$$

(a) Sketch the root locus for $K > 0$ and $K < 0$ when $a = 1/2$.

(b) Repeat part (a) when $a = -1/2$.

(c) With $a = -1/2$, find a value of K for which the closed-loop impulse response is of the form

$$(A + Bn)\alpha^n$$

for some values of the constants A, B, and α, with $|\alpha| < 1$. (*Hint:* What must the denominator of the closed-loop system function look like in this case?)

11.39. Consider the feedback system of Figure P11.39 with

$$H(z) = \frac{1}{1 - \frac{1}{2}z^{-1}}, \quad G(z) = K. \qquad \text{(P11.39–1)}$$

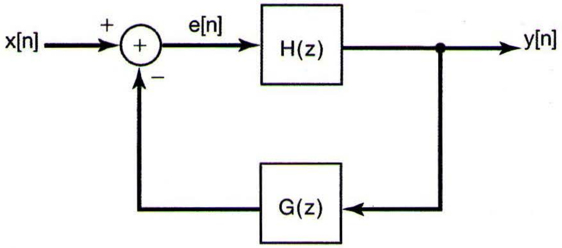


Figure P11.39

(a) Plot the root locus for $K > 0$.

(b) Plot the root locus for $K < 0$. (*Note:* Be careful with this root locus. By applying the angle criterion on the real axis, you will find that as K is decreased from zero, the closed loop approaches $z = +\infty$ along the positive real axis and then returns along the negative real axis from $z = -\infty$. Check that this is in fact the case by explicitly solving for the closed-loop pole as a function of K. At what value of K is the pole at $|z| = \infty$?)

(c) Find the full range of values of K for which the closed-loop system is stable.

(d) The phenomenon observed in part (b) is a direct consequence of the fact that in this example the numerator and denominator of $G(z)H(z)$ have the same degree. When this occurs in a discrete-time feedback system, it means that there is a delay-free loop in the system. That is, the output at a given point in time is being fed back into the system and in turn affects its own value at the same point in time. To see that this is the case in the system we are considering here, write the difference equation relating $y[n]$ and $e[n]$. Then write $e[n]$ in terms of the input and output for the feedback system. Contrast this result with that of the feedback system with

$$H(z) = \frac{1}{1 - \frac{1}{2}z^{-1}}, \quad G(z) = Kz^{-1}. \qquad \text{(P11.39–2)}$$

The primary consequence of having delay-free loops is that such feedback systems cannot be implemented in the form depicted. For example, for the system of eq. (P11.39–1), we cannot first calculate $e[n]$ and then $y[n]$, because $e[n]$ depends on $y[n]$. Note that we *can* perform this type of calculation for the system of eq. (P11.39–2), since $e[n]$ depends on $y[n-1]$.

(e) Show that the feedback system of eq. (P11.39–1) represents a causal system, except for the value of K for which the closed-loop pole is at $|z| = \infty$.

11.40. Consider the discrete-time feedback system depicted in Figure P11.40. The system in the forward path is not very well damped, and we would like to choose the feedback system function so as to improve the overall damping. By using the root-locus method, show that this can be done with

$$G(z) = 1 - \frac{1}{2}z^{-1}.$$

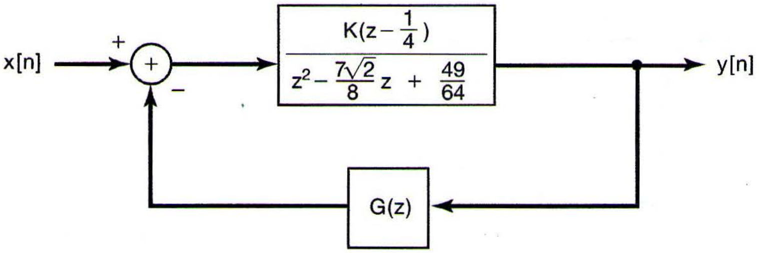

x[n]
+
−
$\frac{K(z-\frac{1}{4})}{z^2-\frac{7\sqrt{2}}{8}z+\frac{49}{64}}$
y[n]
G(z)

Figure P11.40

Specifically, sketch the root locus for $K > 0$, and specify the value of the gain K for which a significant improvement in damping is obtained.

11.41. (a) Consider a feedback system with

$$H(z) = \frac{z+1}{z^2+z+\frac{1}{4}}, \qquad G(z) = \frac{K}{z-1}.$$

(i) Write the closed-loop system function explicitly as a ratio of two polynomials. (The denominator polynomial will have coefficients that depend on K.)

(ii) Show that the sum of the closed-loop poles is independent of K.

(b) More generally, consider a feedback system with system function

$$G(z)H(z) = K\frac{z^m + b_{m-1}z^{m-1} + \cdots + b_0}{z^n + a_{n-1}z^{n-1} + \cdots + a_0}.$$

Show that if $m \leq n-2$, the sum of the closed-loop poles is independent of K.

11.42. Consider again the discrete-time feedback system of Example 11.3:

$$G(z)H(z) = \frac{z}{(z-\frac{1}{2})(z-\frac{1}{4})}.$$

The root loci for $K > 0$ and $K < 0$ are depicted in Figure 11.16.

(a) Consider the root locus for $K > 0$. In this case, the system becomes unstable when one of the closed-loop poles is less than or equal to -1. Find the value of K for which $z = -1$ is a closed-loop pole.

(b) Consider the root locus for $K < 0$. In this case, the system becomes unstable when one of the closed-loop poles is greater than or equal to 1. Find the value of K for which $z = 1$ is a closed-loop pole.

(c) What is the full range of values of K for which the closed-loop system is stable?

11.43. Consider a discrete-time feedback system with

$$G(z)H(z) = \frac{1}{z(z-1)}.$$

(a) Sketch the root locus for $K > 0$ and for $K < 0$.

(b) If you have sketched the root locus correctly for $K > 0$, you will see that the two branches of the root locus cross and exit from the unit circle. Consequently, we can conclude that the closed-loop system is stable for $0 < K < K_0$, where K_0 is the value of the gain for which the two branches intersect the unit circle. At what points on the unit circle do the branches exit from it? What is the value of K_0?

11.44. As mentioned in Section 11.4, the continuous-time Nyquist criterion can be extended to allow for poles of $G(s)H(s)$ on the $j\omega$-axis. In this problem, we will illustrate the general technique for doing this by means of several examples. Consider a continuous-time feedback system with

$$G(s)H(s) = \frac{1}{s(s+1)}. \tag{P11.44–1}$$

When $G(s)H(s)$ has a pole at $s = 0$, we modify the contour of Figure 11.19 by avoiding the origin. To do this, we indent the contour by adding a semicircle of infinitesimal radius ϵ into the right-half plane. [See Figure P11.44(a).] Thus, only a small part of the right-half plane is not enclosed by the modified contour, and its area goes to zero as we let $\epsilon \to 0$. Consequently, as $M \to \infty$, the contour will enclose the entire right-half plane. As in the text, $G(s)H(s)$ is a constant (in this case zero) along the circle of infinite radius. Thus, to plot $G(s)H(s)$ along the contour, we need only plot it for the portion of the contour consisting of the $j\omega$-axis and the infinitesimal circle.

(a) Show that

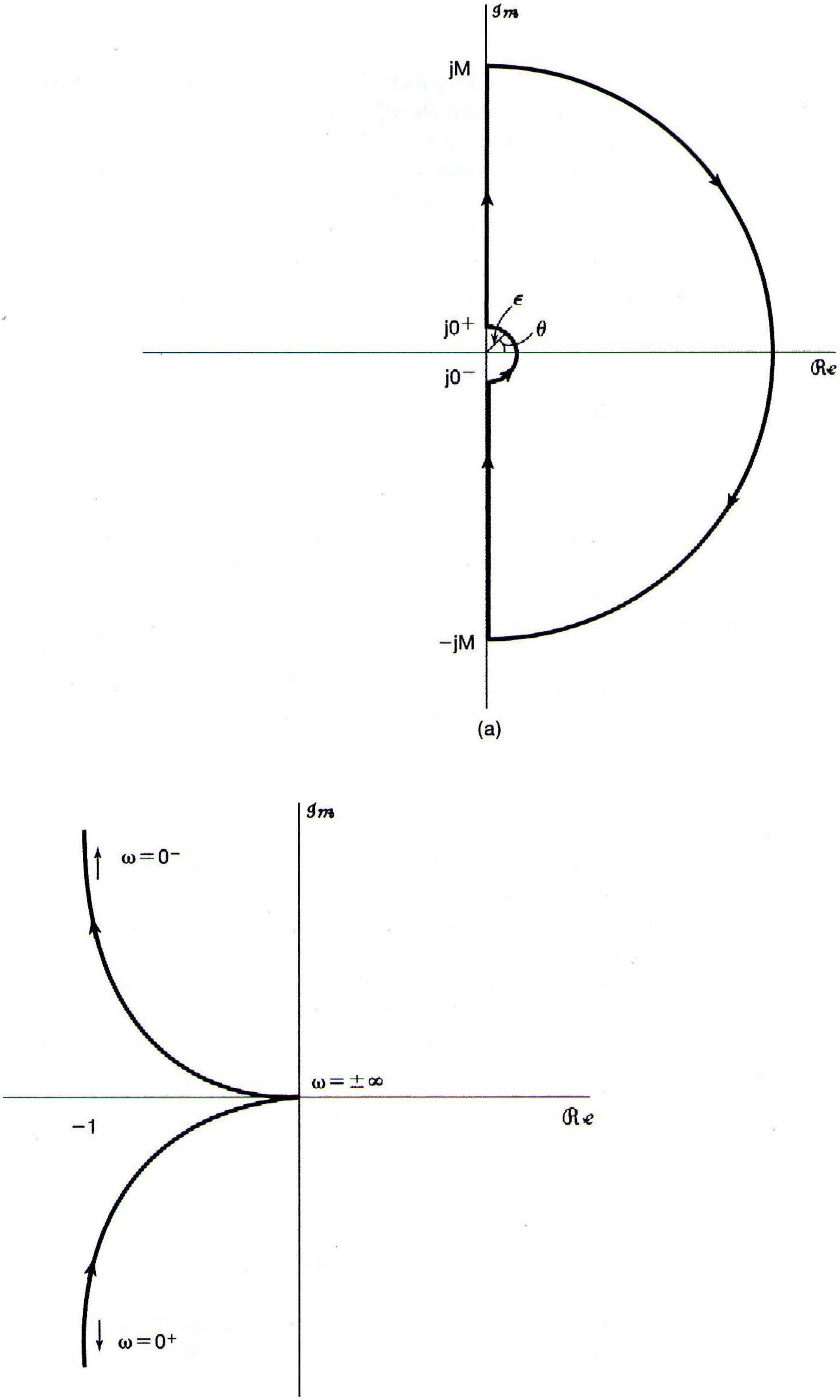
$$\sphericalangle G(j0^+)H(j0^+) = -\frac{\pi}{2}$$

and

$$\sphericalangle G(j0^-)H(j0^-) = \frac{\pi}{2},$$

where $s = j0^-$ is the point where the infinitesimal semicircle meets the $j\omega$-axis just below the origin and $s = j0^+$ is the corresponding point just above the origin.

(b) Use the result of part (a) together with eq. (P11.44–1) to verify that Figure P11.44(b) is an accurate sketch of $G(s)H(s)$ along the portions of the contour


ℐm
jM
ε
j0+
θ
j0−
ℜe
−jM
ℐm
ω = 0−
ω = ±∞
−1
ℜe
ω = 0+

(a)

(b)

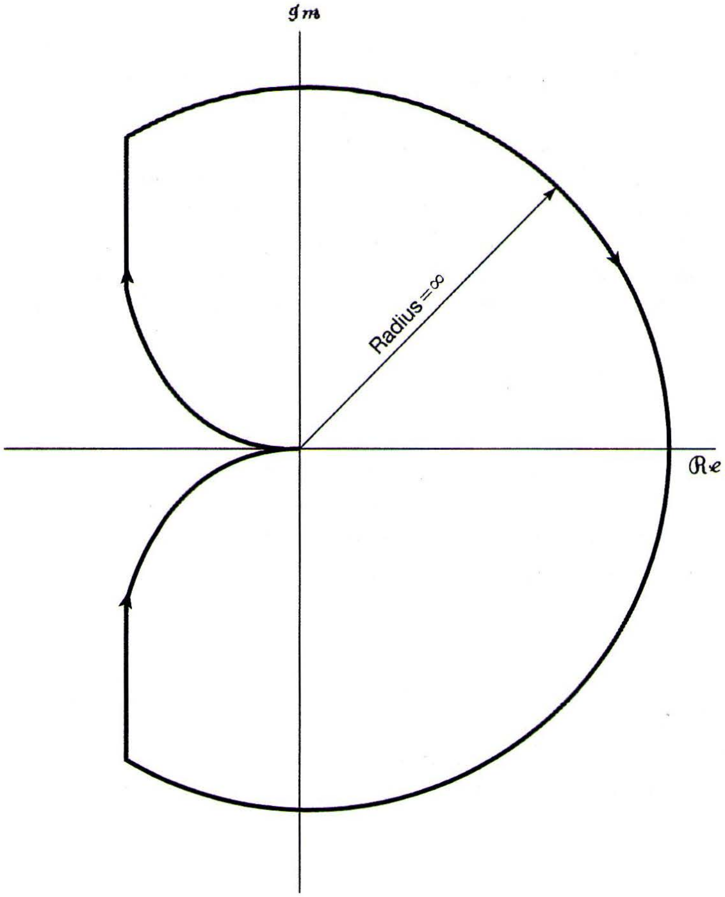
Figure P11.44

from $-j\infty$ to $j0^-$ and $j0^+$ to $j\infty$. In particular, check that $\sphericalangle G(j\omega)H(j\omega)$ and $|G(j\omega)H(j\omega)|$ behave in the manner depicted in the figure.

(c) All that remains to be done is to determine the plot of $G(s)H(s)$ along the small semicircle about $s = 0$. Note that as $\epsilon \to 0$, the magnitude of $G(s)H(s)$ along this contour goes to infinity. Show that as $\epsilon \to 0$, the contribution of the pole at $s = -1$ to $\sphericalangle G(s)H(s)$ along the semicircle is zero. Then show that as $\epsilon \to 0$,

$$\sphericalangle G(s)H(s) = -\theta,$$

where θ is as defined in Figure P11.44(a). Thus, since θ varies from $-\pi/2$ at $s = j0^-$ to $+\pi/2$ at $s = j0^-$ in the counterclockwise direction, $\sphericalangle G(s)H(s)$ must go from $+\pi/2$ at $s = j0^+$ to $-\pi/2$ at $s = j0^+$ in the clockwise direction. The result is the complete Nyquist plot depicted in Figure P11.44(c).



(c)

Figure P11.44 *Continued*

(d) Using the Nyquist plot of Figure P11.44(c), find the range of values of K for which the closed-loop feedback system is stable. (*Note:* As presented in the text, the continuous-time Nyquist criterion states that, for closed-loop system stability, the net number of clockwise encirclements of the point $-1/K$ must equal minus the net number of right-half plane poles of $G(s)H(s)$. In the present example, note that the pole of $G(s)H(s)$ at $s = 0$ is *outside* the modified contour. Consequently, it is *not* included in counting the poles of $G(s)H(s)$ in the right-half plane [i.e., only poles of $G(s)H(s)$ strictly *inside* the right-half plane are counted in applying the Nyquist criterion]. Thus, in this case, since $G(s)H(s)$ has no poles strictly inside the right-half plane, we must have *no* encirclements of the point $s = -1/K$ for closed-loop system stability.)

(e) Follow the steps outlined in parts (a)–(c) to sketch the Nyquist plots for each of the following:

(i) $G(s)H(s) = \frac{(s/10)+1}{s(s+1)}$

(ii) $G(s)H(s) = \frac{1}{s(s+1)^2}$

(iii) $G(s)H(s) = \frac{1}{s^2}$ [be careful in calculating $\sphericalangle G(s)H(s)$ along the infinitesimal semicircle]

(iv) $G(s)H(s) = \frac{s+1}{s(1-s)}$ [be careful in calculating $\sphericalangle G(j\omega)H(j\omega)$ as ω is varied; make sure to take the minus sign in the denominator into account]

(v) $G(s)H(s) = \frac{s+1}{s^2}$ [same remark as for (iii)]

In each case, use the Nyquist criterion to determine the range of values of K (if any such range exists) for which the closed-loop system is stable. Also, use another method (root locus or direct calculation of the closed-loop poles as a function of K) to provide a partial check of the correctness of your Nyquist plot. [*Note:* In sketching the Nyquist plots, you may find it useful to sketch the Bode plots of $G(s)H(s)$ first. It may also be helpful to determine the values of ω for which $G(j\omega)H(j\omega)$ is real.]

(f) Repeat part (e) for:

(i) $G(s)H(s) = \frac{1}{s^2+1}$

(ii) $G(s)H(s) = \frac{s+1}{s^2+1}$

Note: In these cases there are *two* poles on the imaginary axis; accordingly, you will need to modify the contour of Figure 11.19 to avoid each of them. Use infinitesimal semicircles, as in Figure P11.44(a).

11.45. Consider a system with system function

$$H(s) = \frac{1}{(s+1)(s-2)}. \tag{P11.45–1}$$

Because this system is unstable, we would like to devise some method for its stabilization.

(a) Consider first a series compensation scheme as illustrated in Figure P11.45(a). Show that the overall system of this figure is stable if the system function

$$C(s) = \frac{s-2}{s+3}.$$

In practice, this is *not* considered to be a particularly useful way to attempt to stabilize a system. Explain why.

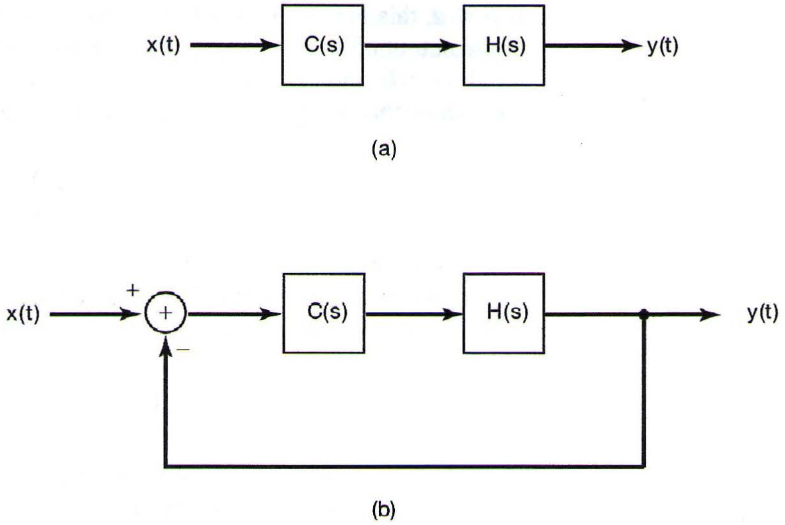


Figure P11.45

(b) Suppose that instead we use a feedback system, as depicted in Figure P11.45(b). Is it possible to stabilize this system using a constant gain, that is,

$$C(s) = K,$$

for the stabilizing element? Justify your answer using Nyquist techniques.

(c) Show that the system of Figure P11.45(b) can be stabilized if $C(s)$ is a proportional plus derivative system—that is, if

$$C(s) = K(s + a).$$

Consider both the case $0 < a < 1$ and the case $a > 1$.

(d) Suppose that

$$C(s) = K(s + 2).$$

Choose the value of K such that the closed-loop system has a pair of complex poles with a damping ratio $\zeta = 1/2$. (*Hint:* In this case, the denominator of the closed-loop system must have the form

$$s^2 + \omega_n s + \omega_n^2$$

for some value of $\omega_n > 0$.)

(e) Pure derivative compensation is both impossible to obtain and undesirable in practice. This is because the required amplification of arbitrarily high frequencies neither can be obtained nor is advisable, as all real systems are subject to some level of high-frequency disturbances. Thus, suppose that we consider a compensator of the form

$$C(s) = K\left(\frac{s + a}{s + b}\right), \quad a, b > 0. \qquad \text{(P11.45–2)}$$

If $b < a$, this is a *lag network*: $\sphericalangle C(j\omega) < 0$ for all $\omega > 0$, so that the phase of the output of the system lags the phase of the input. If $b > a$, $\sphericalangle C(j\omega) > 0$ for all $\omega > 0$, and the system is then called a *lead network.*

(i) Show that it is possible to stabilize the system with the lead compensator

$$C(s) = K\frac{s+\frac{1}{2}}{s+2} \tag{P11.45–3}$$

if K is chosen large enough.

(ii) Show that it is not possible to stabilize the feedback system of Figure P11.45(b) using the lag network

$$C(s) = K\frac{s+3}{s+2}.$$

Hint: Use the results of Problem 11.34 in sketching the root locus. Then determine the points on the $j\omega$-axis that are on the root locus and the values of K for which each of these points is a closed-loop pole. Use this information to prove that for no value of K are *all* of the closed-loop poles in the left-half plane.

11.46. Consider the continuous-time feedback system depicted in Figure P11.46(a).

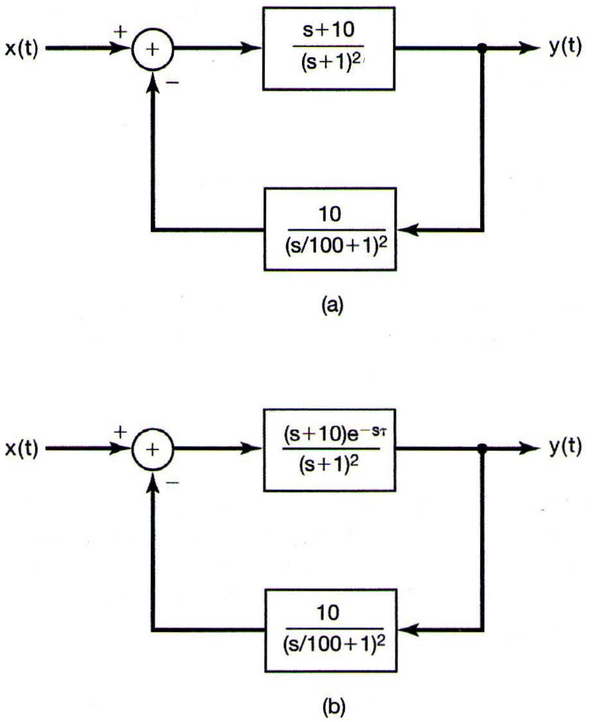


Figure P11.46

(a) Use the straight-line approximations to Bode plots developed in Chapter 6 to obtain a sketch of the log magnitude-phase plot of this system. Estimate the phase and gain margins from your plot.

(b) Suppose that there is an unknown delay within the feedback system, so that the actual feedback system is as shown in Figure P11.46(b). Approximately what is the largest delay τ that can be tolerated before the feedback system becomes unstable? Use your results from part (a) for this calculation.

(c) Calculate more precise values of the phase and gain margins, and compare these to your results in part (a). This should give you some idea of the size of the errors that are incurred in using the approximate Bode plots.

11.47. As mentioned at the end of Section 11.5, the phase and gain margins may provide *sufficient* conditions to ensure that a stable feedback system remains stable. For example, we showed that a stable feedback system will remain stable as the gain is increased, until we reach a limit specified by the gain margin. This does *not* imply (a) that the feedback system cannot be made unstable by *decreasing* the gain or (b) that the system will be unstable for *all* values of gain greater than the gain margin limit. In this problem, we illustrate these two points.

(a) Consider a continuous-time feedback system with

$$G(s)H(s) = \frac{1}{(s-1)(s+2)(s+3)}.$$

Sketch the root locus for this system for $K > 0$. Use the properties of the root locus described in the text and in Problem 11.34 to help you draw the locus accurately. Once you do so, you should see that for small values of the gain K the system is unstable, for larger values of K the system is stable, while for still larger values of K the system again becomes unstable. Find the range of values of K for which the system is stable. *Hint:* Use the same method as is employed in Example 11.2 and Problem 11.35 to determine the values of K at which branches of the root locus pass through the origin and cross the $j\omega$-axis.

If we set our gain somewhere within the stable range that you have just found, we can increase the gain somewhat and maintain stability, but a large enough increase in gain causes the system to become unstable. This maximum amount of increase in gain at which the closed-loop system just becomes unstable is the gain margin. Note that if we *decrease* the gain too much, we can also cause instability.

(b) Consider the feedback system of part (a) with the gain K set at a value of 7. Show that the closed-loop system is stable. Sketch the log magnitude-phase plot of this system, and show that there are two nonnegative values of ω for which $\sphericalangle G(j\omega)H(j\omega) = -\pi$. Further, show that, for one of these values $7|G(j\omega)H(j\omega)| < 1$, and for the other $7|G(j\omega)H(j\omega)| > 1$. The first value provides us with the usual gain margin—that is, the factor $1/|7G(j\omega)H(j\omega)|$ by which we can increase the gain and cause instability. The second provides us with the factor $1/|7G(j\omega)H(j\omega)|$ by which we can decrease the gain and just cause instability.

(c) Consider a feedback system with

$$G(s)H(s) = \frac{(s/100+1)^2}{(s+1)^3}.$$

Sketch the root locus for $K > 0$. Show that two branches of the root locus begin in the left-half plane and, as K is increased, move into the right-half plane and then back into the left-half plane. Do this by examining the equation

$$G(j\omega)H(j\omega) = -\frac{1}{K}.$$

Specifically, by equating the real and imaginary parts of this equation, show that there are two values of $K \geq 0$ for which the closed-loop poles lie on the $j\omega$-axis.

Thus, if we set the gain at a small enough value so that the system is stable, then we can increase the gain up until the point at which the two branches of the root locus intersect the $j\omega$-axis. For a range of values of gain beyond this point, the closed-loop system is unstable. However, if we *continue* to increase the gain, the system will again become stable for K large enough.

(d) Sketch the Nyquist plot for the system of part (c), and confirm the conclusions reached in part (c) by applying the Nyquist criterion. (Make sure to count the *net* number of encirclements of $-1/K$.)

Systems such as that considered in parts (c) and (d) of this problem are often referred to as being *conditionally stable* systems, because their stability properties may change several times as the gain is varied.

11.48. In this problem, we illustrate the discrete-time counterpart of the technique described in Problem 11.44. Specifically, the discrete-time Nyquist criterion can be extended to allow for poles of $G(z)H(z)$ on the unit circle.

Consider a discrete-time feedback system with

$$G(z)H(z) = \frac{z^{-2}}{1-z^{-1}} = \frac{1}{z(z-1)}. \qquad \text{(P11.48–1)}$$

In this case, we modify the contour on which we evaluate $G(z)H(z)$, as illustrated in Figure P11.48(a).

(a) Show that

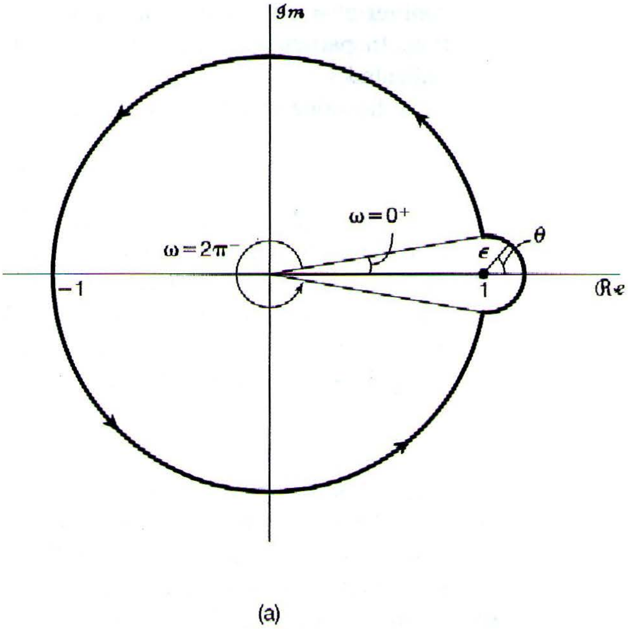
$$\sphericalangle G(e^{j0^+})H(e^{j0^+}) = -\frac{\pi}{2}$$

and

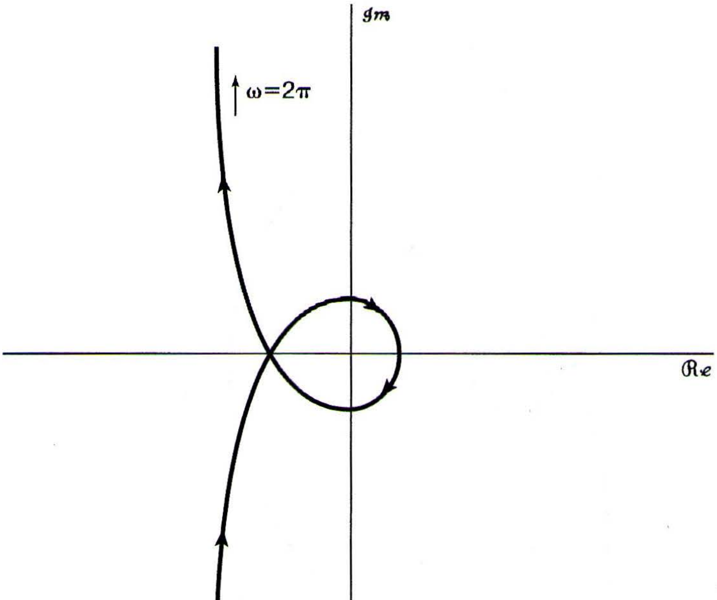
$$\sphericalangle G(e^{j2\pi^-})H(e^{j2\pi^-}) = \frac{\pi}{2},$$

where $z = e^{j2\pi^-}$ is the point below the real axis at which the small semicircle intersects the unit circle and $z = e^{j0^+}$ is the corresponding point above the real axis.

(b) Use the results of part (a) together with eq. (P11.48–1) to verify that Figure P11.48(b) is an accurate sketch of $G(z)H(z)$ along the portion of the



$\mathcal{I}m$
$\omega = 0^+$
$\omega = 2\pi^-$
ϵ
θ
−1
1
$\mathcal{R}e$
(a)



$\mathcal{I}m$
$\omega = 2\pi$
$\omega = 0^+$
$\mathcal{R}e$
(b)

Figure P11.48

contour $z = e^{j\omega}$ as ω varies from 0^+ to $2\pi^-$ in a counterclockwise direction. In particular, verify that the angular variation of $G(e^{j\omega})H(e^{j\omega})$ is as indicated.

(c) Find the value of ω for which $\sphericalangle G(e^{j\omega})H(e^{j\omega}) = -\pi$, and verify that

$$|G(e^{j\omega})H(e^{j\omega})| = 1$$

at this point. [*Hint:* Use the geometrical method for evaluating $\sphericalangle G(e^{j\omega})H(e^{j\omega})$ together with some elementary geometry to determine the value of ω.]

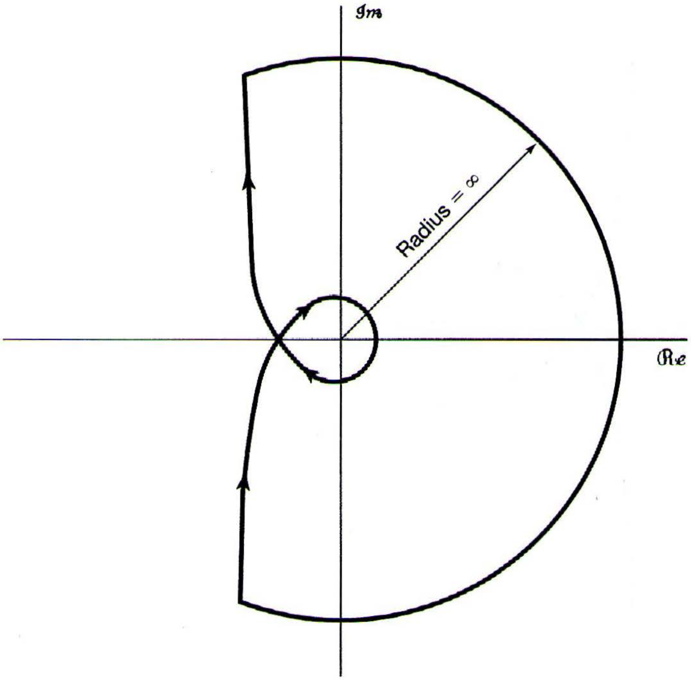
(d) Consider next the plot of $G(z)H(z)$ along the small semicircle about $z = 1$. Note that as $\epsilon \to 0$, the magnitude of $G(z)H(z)$ along this contour goes to infinity. Show that as $\epsilon \to 0$, the contribution of the pole at $z = 0$ to $\sphericalangle G(z)H(z)$ along the semicircle is zero. Then show that as $\epsilon \to 0$,

$$\sphericalangle G(z)H(z) = -\theta,$$

where θ is as defined in Figure P11.48(a).

Thus, since θ varies from $-\pi/2$ to $+\pi/2$ in the counterclockwise direction, $\sphericalangle G(z)H(z)$ varies from $+\pi/2$ to $-\pi/2$ in the clockwise direction. The result is the complete Nyquist plot of Figure P11.48(c).

(e) Using the Nyquist plot, find the range of values of K for which the closed-loop feedback system is stable. [*Note:* Since the pole of $G(z)H(z)$ at $z = 1$ is



(c)

Figure P11.48 *Continued*

inside the modified contour, it is not included in counting the poles of $G(z)H(z)$ outside the unit circle. That is, only poles *strictly outside* the unit circle are counted in applying the Nyquist criterion. Thus, in this case, since $G(z)H(z)$ has no poles strictly outside the unit circle, we must have no encirclements of the point $z = -1/K$ for closed-loop stability.]

(f) Follow the steps outlined in parts (a), (b), and (d) to sketch the Nyquist plots for each of the following:

(i) $\frac{z+\frac{1}{2}+\sqrt{3}}{z-1}$

(ii) $\frac{1}{(z-1)(z+\frac{1}{2}+\sqrt{3})}$

(iii) $\frac{z+1}{z(z-1)}$

(iv) $\frac{z-1/\sqrt{3}}{(z-1)^2}$ [be careful in calculating $\sphericalangle G(z)H(z)$ along the infinitesimal semicircle]

For each of the preceding, use the Nyquist criterion to determine the range of values of K (if any such range exists) for which the closed-loop system is stable. Also, use another method (root locus or direct calculation of the closed-loop poles as a function of K) to provide a partial check of the correctness of your Nyquist plot. *Note:* In sketching the Nyquist plots, you may find it useful to first sketch the magnitude and phase plots as a function of frequency or at least calculate $|G(e^{j\omega})H(e^{j\omega})|$ and $\sphericalangle G(e^{j\omega})H(e^{j\omega})$ at several points. Also, it is helpful to determine the values of ω for which $G(e^{j\omega})H(e^{j\omega})$ is real.

(g) Repeat part (f) for

$$G(z)H(z) = \frac{1}{z^2-1}.$$

In this case there are two poles on the unit circle, and thus, you must modify the contour around each of these by including an infinitesimal semicircle that extends outside the unit circle, thereby placing the pole inside the contour.

EXTENSION PROBLEMS

11.49. In this problem, we provide an illustration of how feedback can be used to increase the bandwidth of an amplifier. Consider an amplifier whose gain falls off at high frequencies. That is, suppose the system function of this amplifier is

$$H(s) = \frac{Ga}{s+a}.$$

(a) What is the dc gain of the amplifier (i.e., the magnitude of its frequency response at 0 frequency)?

(b) What is the system time constant?

(c) Suppose we define the bandwidth of the system as the frequency at which the magnitude of the amplifier frequency response is $1/\sqrt{2}$ times its magnitude at dc. What is the bandwidth of the amplifier?

(d) Suppose we place the amplifier in a feedback loop as depicted in Figure P11.49. What is the dc gain of the closed-loop system? What are the time constant and the bandwidth of the closed-loop system?

(e) Find the value of K that leads to a closed-loop bandwidth that is exactly double the bandwidth of the open-loop amplifier. What are the corresponding closed-loop system time constant and dc gain?

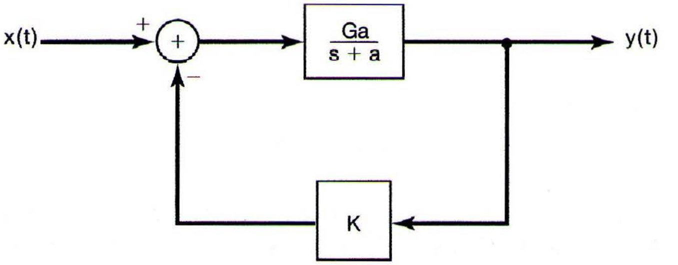


Figure P11.49

11.50. As mentioned in the text, an important class of devices used in the implementation of feedback systems is the class of operational amplifiers. A model for such an amplifier is depicted in Figure P11.50(a). The amplifier's input is the difference between two voltages $v_2(t)$ and $v_1(t)$, and the output voltage is an amplified version of the input; that is,

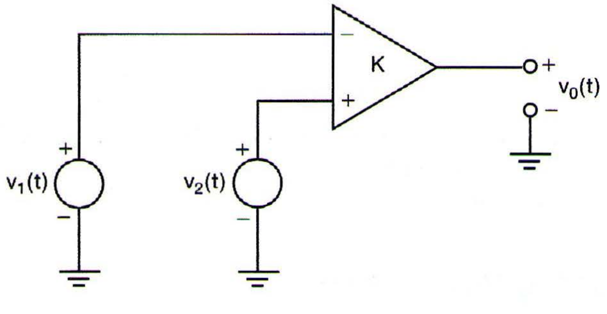


Figure P11.50a

$$v_o(t) = K[v_2(t) - v_1(t)]. \qquad \text{(P11.50–1)}$$

Consider an operational amplifier connection shown in Figure P11.50(b). In this figure, $Z_1(s)$ and $Z_2(s)$ are impedances. (That is, each is the system function of an LTI system whose input is the current flowing through the impedance element and whose output is the voltage across the element.) Making the approximation that the input impedance of the operational amplifier is infinite and that its output impedance is zero, we obtain the following relationship between $V_1(s)$, $V_i(s)$, and $V_o(s)$, the Laplace transforms of $v_1(t)$, $v_i(t)$, and $v_o(t)$, respectively:

$$V_1 = \left[\frac{Z_2(s)}{Z_1(s) + Z_2(s)}\right] V_i(s) + \left[\frac{Z_1(s)}{Z_1(s) + Z_2(s)}\right] V_o(s). \qquad \text{(P11.50–2)}$$

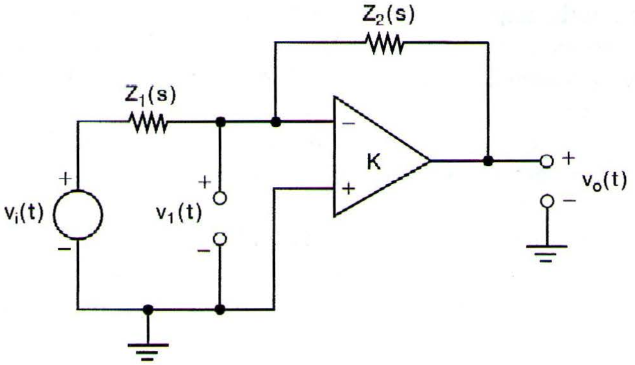


Figure P11.50b

Also, from eq. (P11.50–1) and Figure P11.50(b), we see that

$$V_o(s) = -KV_1(s). \tag{P11.50–3}$$

(a) Show that the system function

$$H(s) = \frac{V_o(s)}{V_i(s)}$$

for the interconnection of Figure P11.50(b) is identical to the overall closed-loop system function for the system of Figure P11.50(c).

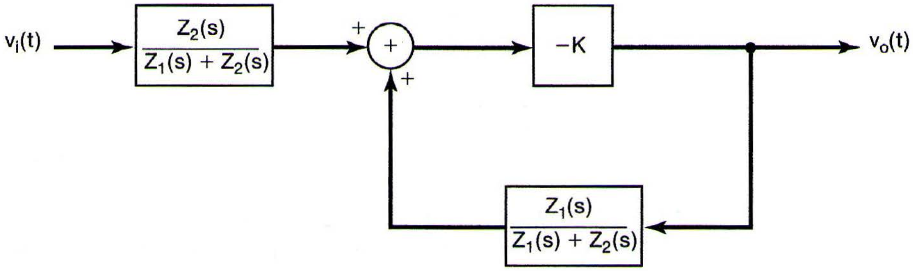

$v_i(t)$
$\frac{Z_2(s)}{Z_1(s) + Z_2(s)}$
+
+
+
−K
$v_o(t)$
$\frac{Z_1(s)}{Z_1(s) + Z_2(s)}$

Figure P11.50c

(b) Show that if $K >> 1$, then

$$H(s) \approx -\frac{Z_2(s)}{Z_1(s)}.$$

11.51. **(a)** Suppose that in Figure P11.50(b) $Z_1(s)$ and $Z_2(s)$ are both pure resistances, say, R_1 and R_2, respectively. A typical value for R_2/R_1 is in the range 1 to 10^3, while a typical value for K is 10^6. Using the results of Problem 11.50(a), calculate the actual system function for this value of K and for R_2/R_1 equal to 1 and then to 10^3, and compare each resulting value to $-R_2/R_1$. This should give you some idea of how good the approximation of Problem 11.50(b) typically is.

(b) One of the important uses of feedback is in the reduction of system sensitivity to variations in parameters. This is particularly important for circuits involving operational amplifiers, which have high gains that may be known only approximately.

(i) Consider the circuit discussed in part (a), with $R_2/R_1 = 10^2$. What is the percentage change in the closed-loop gain of the system if K changes from 10^6 to 5×10^5?

(ii) How large must K be so that a 50% reduction in its value results in only a 1% reduction in the closed-loop gain? Again, take $R_2/R_1 = 10^2$.

11.52. Consider the circuit of Figure P11.52. This circuit is obtained by using

$$Z_1(s) = R, \qquad Z_2(s) = \frac{1}{CS}$$

in Figure P11.50(b). Using the results from Problem 11.50, show that the system behaves approximately like an integrator. In what frequency range (expressed in terms of K, R, and C) does this approximation break down?

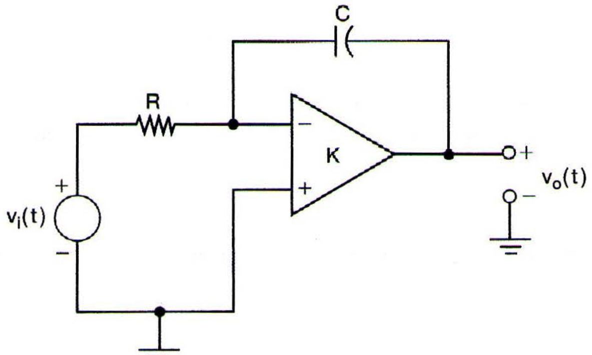


Figure P11.52

11.53. Consider the circuit depicted in Figure P11.53(a), which is obtained from the circuit of Figure P11.50(b) by using $Z_1(s) = R$ and by replacing $Z_2(s)$ with a diode that has an exponential current-voltage relationship. Assume that this relationship is of the form

$$i_d(t) = Me^{qv_d(t)/kT}, \tag{P11.53–1}$$

where M is a constant that depends upon the construction of the diode, q is the charge of an electron, k is Boltzmann's constant, and T is absolute temperature. Note that the idealized relationship of eq. (P11.53–1) assumes that there is no possibility of a negative diode current. Usually, there *is* some small maximum negative value of diode current, but we will neglect this possibility in our analysis.

(a) Assuming that the input impedance of the operational amplifier is infinite and that its output impedance is zero, show that the following relations hold:

$$v_o(t) = v_d(t) + Ri_d(t) + v_i(t), \tag{P11.53–2}$$

$$v_o(t) = -K[v_o(t) - v_d(t)]. \tag{P11.53–3}$$

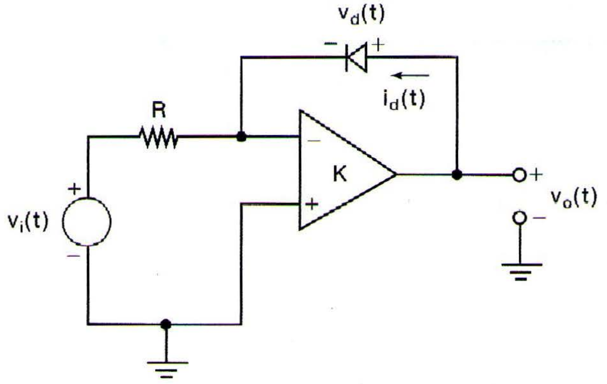


Figure P11.53a

(b) Show that for K large, the relationship between $v_o(t)$ and $v_i(t)$ is essentially the same as in the feedback system of Figure P11.53(b), in which the system in the feedback path is a nonlinear memoryless system with input $v_o(t)$ and output

$$w(t) = RMe^{qv_o(t)/kT}.$$

(c) Show that for K large,

$$v_o(t) \approx \frac{kT}{q} \ln\left(-\frac{v_i(t)}{RM}\right). \tag{P11.53–4}$$

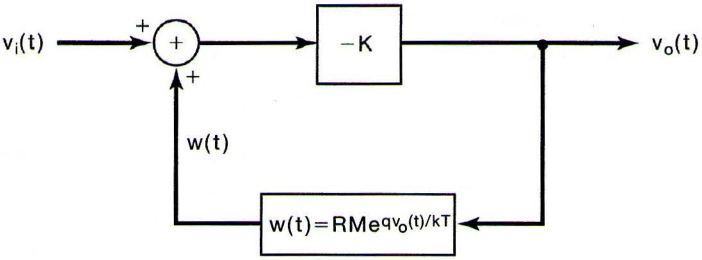


Figure P11.53b

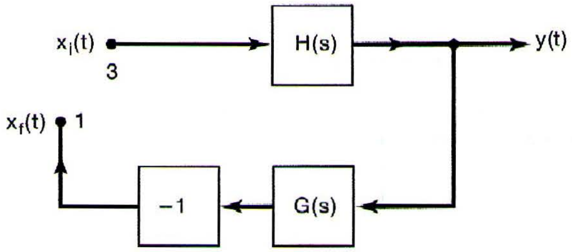
Note that eq. (P11.53–4) makes sense only for a negative $v_i(t)$, which is consistent with the requirement that the diode current cannot be negative. If a positive $v_i(t)$ is applied, the current $i_d(t)$ cannot balance the current through the resistor. Thus, a nonnegligible current is fed into the amplifier, causing it to saturate.

11.54. In this problem, we explore the use of positive feedback for generating oscillating signals.

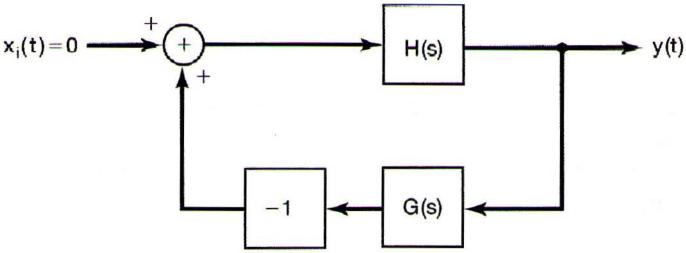
(a) Consider the system illustrated in Figure P11.54(a). Show that $x_f(t) = x_i(t)$ if

$$G(s)H(s) = -1. \tag{P11.54–1}$$

Suppose that we connect terminals 1 and 2 in Figure P11.54(a) and make $x_i(t) = 0$. Then the output of the system should remain unchanged if we



(a)



(b)

Figure P11.54

satisfy eq. (P11.54–1). The system now produces an output without any input. Therefore, the system shown in Figure P11.54(b) is an oscillator, provided that eq. (P11.54–1) is satisfied.

(b) A commonly used oscillator in practice is the sinusoidal oscillator. For such an oscillator, we may rewrite the condition of eq. (P11.54–1) as

$$G(j\omega_0)H(j\omega_0) = -1. \tag{P11.54–2}$$

What is the value of the closed-loop gain for the system shown in Figure P11.54(b) at ω_0 when eq. (P11.54–2) is satisfied?

(c) A sinusoidal oscillator may be constructed on the basis of the principle outlined above by using the circuit shown in Figure P11.54(c). The input to the

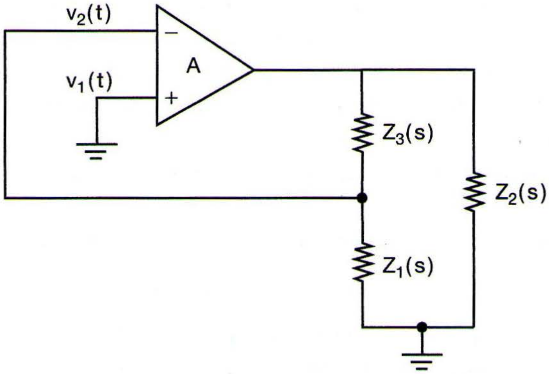


Figure P11.54c

amplifier is the difference between the voltages $v_1(t)$ and $v_2(t)$. In this circuit, the amplifier has a gain of A and an output resistance of R_0. $Z_1(s)$, $Z_2(s)$, and $Z_3(s)$ are impedances. (That is, each is the system function of an LTI system whose input is the current flowing through the impedance element and whose output is the voltage across the element.) It can be shown that, for this circuit,

$$H(s) = \frac{-AZ_L(s)}{Z_L(s) + R_0},$$

where

$$Z_L = \frac{Z_2(s)(Z_1(s) + Z_3(s))}{Z_1(s) + Z_2(s) + Z_3(s)}.$$

Also, we can show that

$$G(s) = \frac{-Z_1(s)}{Z_1(s) + Z_3(s)}.$$

(i) Show that

$$G(s)H(s) = \frac{AZ_1(s)Z_2(s)}{R_0(Z_1(s) + Z_2(s) + Z_3(s)) + Z_2(s)(Z_1(s) + Z_3(s))}.$$

(ii) If $Z_1(s)$, $Z_2(s)$, and $Z_3(s)$ are pure reactances (i.e., inductances or capacitances), we can write $Z_1(j\omega) = jX_1(j\omega)$, $Z_2(j\omega) = jX_2(j\omega)$, and $Z_3(j\omega) = jX_3(j\omega)$, where $X_i(j\omega)$, $i = 1, 2, 3$, are all real. Using the results of parts (b) and (i), show that a necessary condition for the circuit to produce oscillations is

$$X_1(j\omega) + X_2(j\omega) + X_3(j\omega) = 0.$$

(iii) Show also that, in addition to the constraint of part (ii), the constraint $AX_1(j\omega) = X_2(j\omega)$ has to be satisfied for the circuit to produce oscillations. [Since $X_i(j\omega)$ is positive for inductances and negative for capacitances, the latter constraint requires that $Z_1(s)$ and $Z_2(s)$ be reactances of the same type (i.e., both should be inductances or both should be capacitances).]

(iv) Let us assume that $Z_1(s)$ and $Z_2(s)$ are both inductances such that

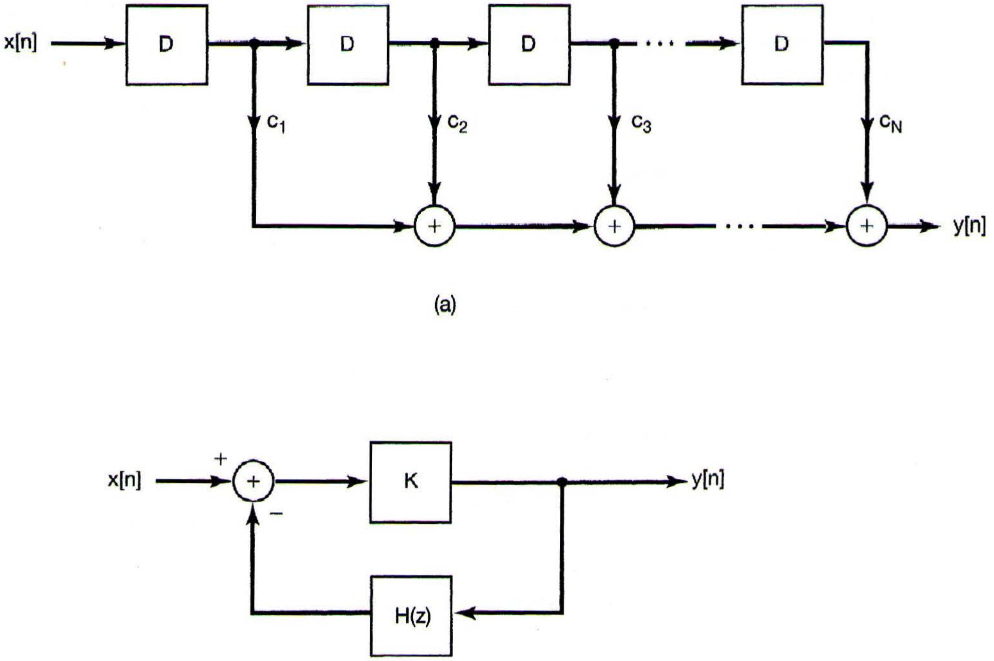
$$X_1(j\omega) = X_2(j\omega) = \omega L.$$

Let us also assume that

$$X_3(j\omega) = -1/(\omega C)$$

is a capacitance. Use the condition derived in (ii) to determine the frequency (in terms of L and C) at which the circuit oscillates.

11.55. **(a)** Consider the nonrecursive discrete-time LTI filter depicted in Figure P11.55(a). Through the use of feedback around this nonrecursive system, a recursive filter can be implemented. To do so, consider the configuration shown in Figure P11.55(b), in which $H(z)$ is the system function of the nonrecursive LTI system of Figure P11.55(a). Determine the overall system function of this feedback system, and find the difference equation relating the input to the output of the overall system.



(a)

(b)

Figure P11.55

(b) Now suppose that $H(z)$ in Figure P11.55(b) is the system function of a recursive LTI system. Specifically, suppose that

$$H(z) = \frac{\sum_{i=1}^{N} c_i z^{-i}}{\sum_{i=1}^{N} d_i z^{-i}}.$$

Show how one can find values of the coefficients $K, c_1, \ldots, c_N$, and $d_0, \ldots, d_N$, such that the closed-loop system function is

$$Q(z) = \frac{\sum_{i=0}^{N} b_i z^{-i}}{\sum_{i=0}^{N} a_i z^{-i}},$$

where the a_i and b_i are specified coefficients.

In this problem, we have seen that the use of feedback provides us with alternative implementations of LTI systems specified by linear constant-coefficient difference equations. The implementation in part (a), consisting of feedback around a nonrecursive system, is particularly interesting, as some technologies are ideally suited to implementing tapped delay-line structures (i.e., systems consisting of chains of delays with taps at each delay whose outputs are weighted and then summed).

11.56. Consider an inverted pendulum mounted on a movable cart, as depicted in Figure P11.56. Here, we have modeled the pendulum as consisting of a massless rod of length L with a mass m attached at the end. The variable $\theta(t)$ denotes the pendulum's angular deflection from the vertical, g is gravitational acceleration, $s(t)$ is the position of the cart with respect to some reference point, $a(t)$ is the acceleration of the cart, and $x(t)$ represents the angular acceleration resulting from any disturbances, such as gusts of wind.

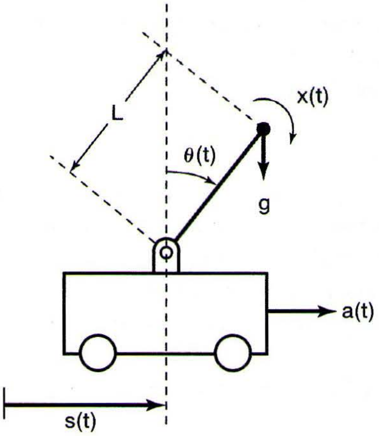


Figure P11.56

Our goal in this problem is to analyze the dynamics of the inverted pendulum and, more specifically, to investigate the problem of balancing the pendulum by a judicious choice of the acceleration $a(t)$ of the cart. The differential equation relating $\theta(t)$, $a(t)$, and $x(t)$ is

$$L\frac{d^2\theta(t)}{dt^2} = g\sin[\theta(t)] - a(t)\cos[\theta(t)] + Lx(t). \qquad \text{(P11.56–1)}$$

This relation merely equates the actual acceleration of the mass along a direction perpendicular to the rod to the applied accelerations [gravity, the disturbance acceleration due to $x(t)$, and the cart's acceleration] along this direction.

Note that eq. (P11.56–1) is a nonlinear differential equation. The detailed, exact analysis of the behavior of the pendulum requires that we examine this equation; however, we can obtain a great deal of insight into the dynamics of the pendulum by performing a linearized analysis. Specifically, let us examine the dynamics of the pendulum when it is nearly vertical [i.e., when $\theta(t)$ is small]. In this case, we can make the approximations

$$\sin[\theta(t)] \approx \theta(t), \cos[\theta(t)] \approx 1. \tag{P11.56–2}$$

(a) Suppose that the cart is stationary [i.e., $a(t) = 0$], and consider the causal LTI system with input $x(t)$ and output $\theta(t)$ described by eq. (P11.56–1), together with the approximations given in eq. (P11.56–2). Find the system function for this system, and show that it has a pole in the right-half of the plane, implying that the system is unstable.

(b) The result of part (a) indicates that if the cart is stationary, any minor angular disturbance caused by $x(t)$ will lead to growing angular deviations from the vertical. Clearly, at some point, these deviations will become sufficiently large so that the approximations of eq. (P11.56–2) will no longer be valid. At this point the linearized analysis is no longer accurate, but the fact that it is accurate for small angular displacements allows us to conclude that the vertical equilibrium position is unstable, since small angular displacements will grow rather than diminish.

We now wish to consider the problem of stabilizing the vertical position of the pendulum by moving the cart in an appropriate fashion. Suppose we try proportional feedback—that is,

$$a(t) = K\theta(t).$$

Assume that $\theta(t)$ is small, so that the approximations in eq. (P11.56–2) are valid. Draw a block diagram of the linearized system with $\theta(t)$ as the output, $x(t)$ as the external input, and $a(t)$ as the signal that is fed back. Show that the resulting closed-loop system is unstable. Find a value of K such that if $x(t) = \delta(t)$, the pendulum will sway back and forth in an undamped oscillatory fashion.

(c) Consider using the proportional-plus-derivative (PD) feedback,

$$a(t) = K_1\theta(t) + K_2\frac{d\theta(t)}{dt}.$$

Show that one can find values of K_1 and K_2 that stabilize the pendulum. In fact, using

$$g = 9.8 \text{ m/sec}^2$$

and (P11.56–3)

$$L = 0.5 \text{ m},$$

choose values of K_1 and K_2 so that the damping ratio of the closed loop system is 1 and the natural frequency is 3 rad/sec.

11.57. In this problem, we consider several examples of the design of tracking systems. Consider the system depicted in Figure P11.57. Here, $H_p(s)$ is the system whose output is to be controlled, and $H_c(s)$ is the compensator to be designed. Our objective in choosing $H_c(s)$ is that we would like the output $y(t)$ to follow the input $x(t)$. In particular, in addition to stabilizing the system, we would also like to design the system so that the error $e(t)$ decays to zero for certain specified inputs.

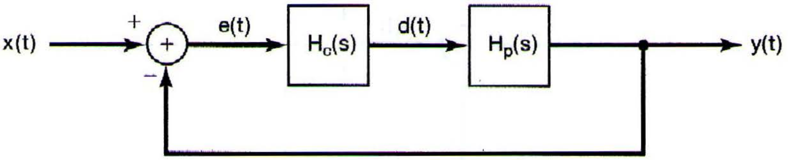


Figure P11.57

(a) Suppose that

$$H_p(s) = \frac{\alpha}{s+\alpha}, \quad \alpha \neq 0. \tag{P11.57–1}$$

Show that if $H_c(s) = K$ (which is known as *proportional* or *P* control), we can choose K so as to stabilize the system and so that $e(t) \to 0$ if $x(t) = \delta(t)$. Show that we *cannot* get $e(t) \to 0$ if $x(t) = u(t)$.

(b) Again let $H_p(s)$ be as in eq. (P11.57–1), and suppose that we use *proportional-plus-integral* (PI) control—that is,

$$H_c(s) = K_1 + \frac{K_2}{s}.$$

Show that we can choose K_1 and K_2 so as to stabilize the system, and we can also get $e(t) \to 0$ if $x(t) = u(t)$. Thus, the system can track a step. In fact, this illustrates a basic and important principle in feedback system design: To track a step $[X(s) = 1/s]$, we need an integrator $(1/s)$ in the feedback system. An extension of this principle is considered in the next problem.

(c) Suppose that

$$H_p(s) = \frac{1}{(s-1)^2}.$$

Show that we cannot stabilize this system with a PI controller, but that we can stabilize it and have it track a step if we use *proportional-plus-integral-plus-differential* (PID) control, i.e.,

$$H_c(s) = K_1 + \frac{K_2}{s} + K_3 s.$$

11.58. In Problem 11.57, we discussed how the presence of an integrator in a feedback system can make it possible for the system to track a step input with zero error in the steady state. In this problem, we extend the idea. Consider the feedback system depicted in Figure P11.58, and suppose that the overall closed-loop system is stable. Suppose also that

$$H(s) = \frac{K\prod_{k=1}^{m}(s-\beta_k)}{s^l\prod_{k=1}^{n-l}(s-\alpha_k)},$$

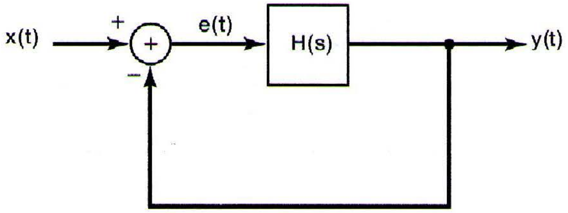


Figure P11.58

where the α_k and β_k are given nonzero numbers and l is a positive integer. The feedback system of Figure P11.58 is often referred to as a *Type l* feedback system.

(a) Use the final-value theorem (Section 9.5.10) to show that a Type 1 feedback system can track a step—that is, that

$$e(t) \to 0 \qquad \text{if} \quad x(t) = u(t).$$

(b) Similarly, show that a Type 1 system cannot track a ramp, but rather, that

$$e(t) \to \text{a finite constant if} \qquad x(t) = u_{-2}(t).$$

(c) Show that, for a Type 1 system, unbounded results ensue if

$$x(t) = u_{-k}(t)$$

with $k > 2$.

(d) More generally, show that, for a Type l system:
(i) $e(t) \to 0$ if $x(t) = u_{-k}(t)$ with $k \leq l$
(ii) $e(t) \to$ a finite constant if $x(t) = u_{(-l+1)}(t)$
(iii) $e(t) \to \infty$ if $x(t) = u_{-k}(t)$ with $k > l+1$

11.59. **(a)** Consider the discrete-time feedback system of Figure P11.59. Suppose that

$$H(z) = \frac{1}{(z-1)(z+\frac{1}{2})}.$$

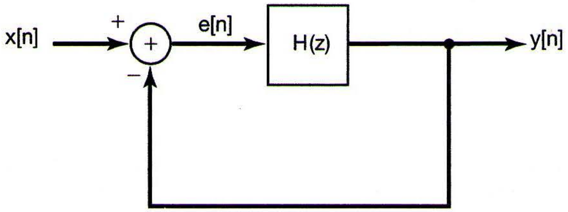


Figure P11.59

Show that this system can track a unit step in the sense that if $x[n] = u[n]$, then

$$\lim_{n\to\infty} e[n] = 0. \qquad \text{(P11.59–1)}$$

(b) More generally, consider the feedback system of Figure P11.59, and assume that the closed-loop system is stable. Suppose that $H(z)$ has a pole at $z = 1$.

Show that the system can track a unit step. [*Hint:* Express the transform $E(z)$ of $e[n]$ in terms of $H(z)$ and the transform of $u[n]$; explain why all the poles of $E(z)$ are inside the unit circle.]

(c) The results of parts (a) and (b) are discrete-time counterparts of the results for continuous-time systems discussed in Problems 11.57 and 11.58. In discrete time, we can also consider the design of the systems that track specified inputs *perfectly* after a finite number of steps. Such systems are known as *deadbeat feedback systems.*

Consider the discrete-time system of Figure P11.59 with

$$H(z) = \frac{z^{-1}}{1 - z^{-1}}.$$

Show that the overall closed-loop system is a deadbeat feedback system with the property that it tracks a step input exactly after one step: that is, if $x[n] = u[n]$, then $e[n] = 0,\ n \geq 1$.

(d) Show that the feedback system of Figure P11.59 with

$$H(z) = \frac{\frac{3}{4}z^{-1} + \frac{1}{4}z^{-2}}{(1 + \frac{1}{4}z^{-1})(1 - z^{-1})}$$

is a deadbeat system with the property that the output tracks a unit step perfectly after a finite number of steps. At what time step does the error $e[n]$ first settle to zero?

(e) More generally, for the feedback system of Figure P11.59, find $H(z)$ so that $y[n]$ perfectly tracks a unit step for $n \geq N$ and, in fact, so that

$$e[n] = \sum_{k=0}^{N-1} a_k \delta[n - k], \qquad \text{(P11.59–2)}$$

where the a_k are specified constants: Hint: Use the relationship between $H(z)$ and $E(z)$ when the input is a unit step and $e[n]$ is given by eq. (P11.59–2).

(f) Consider the system of Figure P11.59 with

$$H(z) = \frac{z^{-1} + z^{-2} - z^{-3}}{(1 + z^{-1})(1 - z^{-1})^2}.$$

Show that this system tracks a ramp $x[n] = (n + 1)u[n]$ exactly after two time steps.

11.60. In this problem, we investigate some of the properties of sampled-data feedback systems and illustrate the use of such systems. Recall from Section 11.2.4 that in a sampled-data feedback system the output of a continuous-time system is sampled. The resulting sequence of samples is processed by a discrete-time system, the output of which is converted to a continuous-time signal that in turn is fed back and subtracted from the external input to produce the actual input to the continuous-time system.

(a) Consider the system within dashed lines in Figure 11.6(b). This is a discrete-time system with input $e[n]$ and output $p[n]$. Show that it is an LTI system. As we have indicated in the figure, we will let $F(z)$ denote the system function of this system.

(b) Show that in Figure 11.6(b) the discrete-time system with system function $F(z)$ is related to the continuous-time system with system function $H(s)$ by means of a *step-invariant* transformation. That is, if $s(t)$ is the step response of the continuous-time system and $q[n]$ is the step response of the discrete-time system, then

$$q[n] = s(nT) \qquad \text{for all } n.$$

(c) Suppose that

$$H(s) = \frac{1}{s-1}, \ \mathcal{R}e\{s\} > 1.$$

Show that

$$F(z) = \frac{(e^T - 1)z^{-1}}{1 - e^T z^{-1}}, \qquad |z| > e^T.$$

(d) Suppose that $H(s)$ is as in part (c) and that $G(z) = K$. Find the range of values of K for which the closed-loop discrete-time system of Figure 11.6(b) is stable.

(e) Suppose that

$$G(z) = \frac{K}{1 + \frac{1}{2}z^{-1}}.$$

Under what conditions on T can we find a value of K that stabilizes the overall system? Find a particular pair of values for K and T that yield a stable closed-loop system. *Hint:* Examine the root locus, and find the values for which the poles enter or leave the unit circle.

APPENDIX
PARTIAL-FRACTION EXPANSION

A.1 INTRODUCTION

The purpose of this appendix is to describe the technique of partial-fraction expansion. This tool is of great value in the study of signals and systems; in particular, it is very useful in inverting Fourier, Laplace, or z-transforms and in analyzing LTI systems described by linear constant-coefficient differential or difference equations. The method of partial-fraction expansion consists of taking a function that is the ratio of polynomials and expanding it as a linear combination of simpler terms of the same type. The determination of the coefficients in the linear combination is the basic problem to be solved in obtaining the expansion. As we will see, this is a relatively straightforward problem in algebra that can be solved very efficiently with a bit of "bookkeeping."

To illustrate the basic idea behind and role of partial-fraction expansion, consider the analysis developed in Section 6.5.2 for a second-order continuous-time LTI system specified by the differential equation

$$\frac{d^2y(t)}{dt^2} + 2\zeta\omega_n\frac{dy(t)}{dt} + \omega_n^2 y(t) = \omega_n^2 x(t). \tag{A.1}$$

The frequency response of this system is

$$H(j\omega) = \frac{\omega_n^2}{(j\omega)^2 + 2\zeta\omega_n(j\omega) + \omega_n^2}, \tag{A.2}$$

or, if we factor the denominator,

$$H(j\omega) = \frac{\omega_n^2}{(j\omega - c_1)(j\omega - c_2)}, \tag{A.3}$$

where

$$\begin{aligned} c_1 &= -\zeta\omega_n + \omega_n\sqrt{\zeta^2 - 1}, \\ c_2 &= -\zeta\omega_n - \omega_n\sqrt{\zeta^2 - 1}. \end{aligned} \tag{A.4}$$

Having $H(j\omega)$, we are in a position to answer a variety of questions related to the system. For example, to determine the impulse response of the system, recall that for any number α with $\mathcal{R}e\{s\} < 0$, the Fourier transform of

$$x_1(t) = e^{\alpha t}u(t) \tag{A.5}$$

is

$$X_1(j\omega) = \frac{1}{j\omega - \alpha}, \tag{A.6}$$

while if

$$x_2(t) = te^{\alpha t}u(t), \tag{A.7}$$

then

$$X_2(j\omega) = \frac{1}{(j\omega - \alpha)^2}. \tag{A.8}$$

Therefore, if we can expand $H(j\omega)$ as a sum of terms of the form of eq. (A.6) or (A.8), we can determine the inverse transform of $H(j\omega)$ by inspection. For example, in Section 6.5.2 we noted that when $c_1 \neq c_2$, $H(j\omega)$ in eq. (A.3) could be rewritten in the form

$$H(j\omega) = \left(\frac{\omega_n^2}{c_1 - c_2}\right)\frac{1}{j\omega - c_1} + \left(\frac{\omega_n^2}{c_2 - c_1}\right)\frac{1}{j\omega - c_2}. \tag{A.9}$$

In this case, the Fourier transform pair of eqs. (A.5) and (A.6) allows us to write down immediately the inverse transform of $H(j\omega)$ as

$$h(t) = \left[\frac{\omega_n^2}{c_1 - c_2}e^{c_1 t} + \frac{\omega_n^2}{c_2 - c_1}e^{c_2 t}\right]u(t). \tag{A.10}$$

While we have phrased the preceding discussion in terms of continuous-time Fourier transforms, similar concepts also arise in discrete-time Fourier analysis and in the use of Laplace and z-transforms. In all of these cases, we encounter the important class of *rational transforms*—that is, transforms that are ratios of polynomials in some variable. Also, in each of these contexts, we find reasons for expanding these transforms as sums of simpler terms such as in eq. (A.9). In this section, in order to develop a general procedure for calculating the expansions, we consider rational functions of a general variable v; that is, we examine functions of the form

$$H(v) = \frac{\beta_m v^m + \beta_{m-1}v^{m-1} + \ldots + \beta_1 v + \beta_0}{\alpha_n v^n + \alpha_{n-1}v^{n-1} + \ldots + \alpha_1 v + \alpha_0}. \tag{A.11}$$

For continuous-time Fourier analysis $(j\omega)$ plays the role of v, while for Laplace transforms that role is played by the complex variable s. In discrete-time Fourier analysis, v is usually taken to be $e^{-j\omega}$, while for z-transforms, we can use either z^{-1} or z. After we have developed the basic techniques of partial-fraction expansion, we will illustrate their application to the analysis of both continuous-time and discrete-time LTI systems.

A.2 PARTIAL-FRACTION EXPANSION AND CONTINOUS-TIME SIGNALS AND SYSTEMS

For our purposes, it is convenient to consider rational functions in one of two standard forms. The second of these, which is often useful in the analysis of discrete-time signals and systems, will be discussed shortly. The first of the standard forms is

$$G(v) = \frac{b_{n-1}v^{n-1} + b_{n-2}v^{n-2} + \ldots + b_1 v + b_0}{v^n + a_{n-1}v^{n-1} + \ldots + a_1 v + a_0}. \tag{A.12}$$

In this form the coefficient of the highest order term in the denominator is 1, and the order of the numerator is at least one less than the order of the denominator. (The order of the numerator will be less than $n-1$ if $b_{n-1}=0$.)

If we are given $H(v)$ in the form of eq. (A.11), we can obtain a rational function of the form of eq. (A.12) by performing two straightforward calculations. First, we divide both the numerator and the denominator of $H(v)$ by α_n. This yields

$$H(v) = \frac{\gamma_m v^m + \gamma_{m-1} v^{m-1} + \ldots + \gamma_1 v + \gamma_0}{v^n + a_{n-1} v^{n-1} + \ldots + a_1 v + a_0}, \tag{A.13}$$

where

$$\gamma_m = \frac{\beta_m}{\alpha_n}, \qquad \gamma_{m-1} = \frac{\beta_{m-1}}{\alpha_n}, \qquad \ldots,$$
$$a_{n-1} = \frac{\alpha_{n-1}}{\alpha_n}, \qquad a_{n-2} = \frac{\alpha_{n-2}}{\alpha_n}, \qquad \ldots.$$

If $m < n$, $H(v)$ is called a *strictly proper* rational function, and in this case, letting $b_0 = \gamma_0, b_1 = \gamma_1, \ldots, b_m = \gamma_m$, and setting any remaining b's equal to zero, we see that $H(v)$ in eq. (A.13) is already of the form of eq. (A.12). In most of the discussions in this book in which rational functions are considered, we are concerned primarily with strictly proper rational functions. However, if $H(v)$ is not proper (i.e., if $m \geq n$), we can perform a preliminary calculation that allows us to write $H(v)$ as the sum of a polynomial in v and a strictly proper rational function. That is,

$$\begin{aligned} H(v) = {} & c_{m-n} v^{m-n} + c_{m-n-1} v^{m-n-1} + \ldots + c_1 v + c_0 \\ & + \frac{b_{n-1} v^{n-1} + b_{n-2} v^{n-2} + \ldots + b_1 v + b_0}{v^n + a_{n-1} v^{n-1} + \ldots + a_1 v + a_0}. \end{aligned} \tag{A.14}$$

The coefficients $c_0, c_1, \ldots, c_{m-n}$ and $b_0, b_1, \ldots, b_{n-1}$ can be obtained by equating eqs. (A.13) and (A.14) and then multiplying through by the denominator. This yields

$$\begin{aligned} \gamma_m v^m + \ldots + \gamma_1 v + \gamma_0 = {} & b_{n-1} v^{n-1} + \ldots + b_1 v + b_0 \\ & + (c_{m-n} v^{m-n} + \ldots + c_0)(v^n + a_{n-1} v^{n-1} + \ldots + a_0). \end{aligned} \tag{A.15}$$

By equating the coefficients of equal powers of v on both sides of eq. (A.15), we can determine the c's and b's in terms of the a's and γ's. For example, if $m = 2$ and $n = 1$, so that

$$H(v) = \frac{\gamma_2 v^2 + \gamma_1 v + \gamma_0}{v + a_1} = c_1 v + c_0 + \frac{b_0}{v + a_1}, \tag{A.16}$$

then eq. (A.15) becomes

$$\begin{aligned} \gamma_2 v^2 + \gamma_1 v + \gamma_0 &= b_0 + (c_1 v + c_0)(v + a_1) \\ &= b_0 + c_1 v^2 + (c_0 + a_1 c_1) v + a_1 c_0. \end{aligned}$$

Equating the coefficients of equal powers of v, we obtain the equations

$$\begin{aligned} \gamma_2 &= c_1, \\ \gamma_1 &= c_0 + a_1 c_1, \\ \gamma_0 &= b_0 + a_1 c_0. \end{aligned}$$

The first equation yields the value of c_1, which can then be used in the second to solve for c_0, which in turn can be used in the third to solve for b_0. The result is

$$\begin{aligned} c_1 &= \gamma_2, \\ c_0 &= \gamma_1 - a_1\gamma_2, \\ b_0 &= \gamma_0 - a_1(\gamma_1 - a_1\gamma_2). \end{aligned}$$

The general case of eq. (A.15) can be solved in an analogous fashion.

Our goal now is to focus on the proper rational function $G(v)$ in eq. (A.12) and to expand it into a sum of simpler proper rational functions. To see how this can be done, consider the case of $n = 3$, so that eq. (A.12) reduces to

$$G(v) = \frac{b_2v^2 + b_1v + b_0}{v^3 + a_2v^2 + a_1v + a_0}. \tag{A.17}$$

As a first step, we factor the denominator of $G(v)$ in order to write it in the form

$$G(v) = \frac{b_2v^2 + b_1v + b_0}{(v - \rho_1)(v - \rho_2)(v - \rho_3)}. \tag{A.18}$$

Assuming for the moment that the roots ρ_1, ρ_2, and ρ_3 of the denominator are all distinct, we would like to expand $G(v)$ into a sum of the form

$$G(v) = \frac{A_1}{v - \rho_1} + \frac{A_2}{v - \rho_2} + \frac{A_3}{v - \rho_3}. \tag{A.19}$$

The problem, then, is to determine the constants A_1, A_2, and A_3. One approach is to equate eqs. (A.18) and (A.19) and to multiply through the denominator. In this case, we obtain the equation

$$\begin{aligned} b_2v^2 + b_1v + b_0 = {} & A_1(v - \rho_2)(v - \rho_3) \\ & + A_2(v - \rho_1)(v - \rho_3) \\ & + A_3(v - \rho_1)(v - \rho_2). \end{aligned} \tag{A.20}$$

By expanding the right-hand side of eq. (A.20) and then equating coefficients of equal powers of v, we obtain a set of linear equations that can be solved for A_1, A_2, and A_3.

Although this approach always works, there is a much easier method. Consider eq. (A.19), and suppose that we would like to calculate A_1. Then, multiplying through by $v - \rho_1$, we obtain

$$(v - \rho_1)G(v) = A_1 + \frac{A_2(v - \rho_1)}{v - \rho_2} + \frac{A_3(v - \rho_1)}{v - \rho_3}. \tag{A.21}$$

Since ρ_1, ρ_2, and ρ_3 are distinct, the last two terms on the right-hand side of eq. (A.21) are zero for $v = \rho_1$. Therefore,

$$A_1 = [(v - \rho_1)G(v)]\big|_{v=\rho_1}, \tag{A.22}$$

or, using eq. (A.18),

$$A_1 = \frac{b_2\rho_1^2 + b_1\rho_1 + b_0}{(\rho_1 - \rho_2)(\rho_1 - \rho_3)}. \tag{A.23}$$

Similarly,

$$A_2 = [(v - \rho_2)G(v)]|_{v=\rho_2} = \frac{b_2\rho_2^2 + b_1\rho_2 + b_0}{(\rho_2 - \rho_1)(\rho_2 - \rho_3)}, \tag{A.24}$$

$$A_3 = [(v - \rho_3)G(v)]|_{v=\rho_3} = \frac{b_2\rho_3^2 + b_1\rho_3 + b_0}{(\rho_3 - \rho_1)(\rho_3 - \rho_2)}. \tag{A.25}$$

Suppose now that $\rho_1 = \rho_3 \neq \rho_2$; that is,

$$G(v) = \frac{b_2v^2 + b_1v + b_0}{(v - \rho_1)^2(v - \rho_2)}. \tag{A.26}$$

In this case, we look for an expansion of the form

$$G(v) = \frac{A_{11}}{v - \rho_1} + \frac{A_{12}}{(v - \rho_1)^2} + \frac{A_{21}}{v - \rho_2}. \tag{A.27}$$

Here, we need the $1/(v - \rho_1)^2$ term in order to obtain the correct denominator in eq. (A.26) when we collect terms over a least common denominator. We also need to include the $1/(v - \rho_1)$ term in general. To see why this is so, consider equating eqs. (A.26) and (A.27) and multiplying them through by the denominator of eq. (A.26):

$$\begin{aligned} b_2v^2 + b_1v + b_0 &= A_{11}(v - \rho_1)(v - \rho_2) \\ &\quad + A_{12}(v - \rho_2) + A_{21}(v - \rho_1)^2. \end{aligned} \tag{A.28}$$

Again, if we equate coefficients of equal powers of v, we obtain three equations (for the coefficients of the v^0, v^1, and v^2 terms). If we omit the A_{11} term in eq. (A.27), we will then have three equations in two unknowns, which in general will not have a solution. By including this term, we can always find a solution. In this case also, however, there is a much simpler method. Consider eq. (A.27) and multiply through by $(v - \rho_1)^2$:

$$(v - \rho_1)^2G(v) = A_{11}(v - \rho_1) + A_{12} + \frac{A_{21}(v - \rho_1)^2}{v - \rho_2}. \tag{A.29}$$

From the preceding example, we see immediately how to determine A_{12}:

$$A_{12} = [(v - \rho_1)^2G(v)]|_{v=\rho_1} = \frac{b_2\rho_1^2 + b_1\rho_1 + b_0}{\rho_1 - \rho_2}. \tag{A.30}$$

As for A_{11}, suppose that we differentiate eq. (A.29) with respect to v:

$$\frac{d}{dv}[(v - \rho_1)^2G(v)] = A_{11} + A_{21}\left[\frac{2(v - \rho_1)}{v - \rho_2} - \frac{(v - \rho_1)^2}{(v - \rho_2)^2}\right]. \tag{A.31}$$

It is then apparent that the final term in eq. (A.31) is zero for $v = \rho_1$, and therefore,

$$\begin{aligned} A_{11} &= \left[\frac{d}{dv}(v - \rho_1)^2G(v)\right]\Bigg|_{v=\rho_1} \\ &= \frac{2b_2\rho_1 + b_1}{\rho_1 - \rho_2} - \frac{b_2\rho_1^2 + b_1\rho_1 + b_0}{(\rho_1 - \rho_2)^2}. \end{aligned} \tag{A.32}$$

Finally, by multiplying eq. (A.27) by $v - \rho_2$, we find that

$$A_{21} = [(v - \rho_2)G(v)]\big|_{v=\rho_2} = \frac{b_2\rho_2^2 + b_1\rho_2 + b_0}{(\rho_2 - \rho_1)^2}. \tag{A.33}$$

This example illustrates all of the basic ideas behind partial-fraction expansion in the general case. Specifically, suppose that the denominator of $G(v)$ in eq. (A.12) has distinct roots $\rho_1, \ldots, \rho_\tau$ with *multiplicities* $\sigma_1, \ldots, \sigma_\tau$; that is,

$$G(v) = \frac{b_{n-1}v^{n-1} + \ldots + b_1 v + b_0}{(v - \rho_1)^{\sigma_1}(v - \rho_2)^{\sigma_2}\ldots(v - \rho_r)^{\sigma_r}}. \tag{A.34}$$

In this case, $G(v)$ has a partial-fraction expansion of the form

$$\begin{aligned} G(v) &= \frac{A_{11}}{v - \rho_1} + \frac{A_{12}}{(v - \rho_1)^2} + \ldots + \frac{A_{1\sigma_1}}{(v - \rho_1)^{\sigma_1}} \\ &\quad + \frac{A_{21}}{v - \rho_2} + \ldots + \frac{A_{2\sigma_2}}{(v - \rho_2)^{\sigma_2}} \\ &\quad + \ldots + \frac{A_{r1}}{v - \rho_r} + \ldots + \ldots + \frac{A_{r\sigma_r}}{(v - \rho_r)^{\sigma_r}} \\ &= \sum_{i=1}^{r}\sum_{k=1}^{\sigma_i} \frac{A_{ik}}{(v - \rho_i)^k}, \end{aligned} \tag{A.35}$$

where the A_{ik} are computed from the equation[1]

$$\boxed{A_{ik} = \frac{1}{(\sigma_i - k)!}\left[\frac{d^{\sigma_i - k}}{dv^{\sigma_i - k}}[(v - \rho_i)^{\sigma_i}G(v)]\right]\Bigg|_{v=\rho_i}.} \tag{A.36}$$

This result can be checked much as in the example: Multiply both sides of eq. (A.35) by $(v - \rho_i)^{\sigma_i}$ and differentiate repeatedly, until A_{ik} is no longer multiplied by a power of $v - \rho_i$. Then set $v = \rho_i$.

Example A.1

In Example 4.25, we examine an LTI system described by the differential equation

$$\frac{d^2y(t)}{dt^2} + 4\frac{dy(t)}{dt} + 3y(t) = \frac{dx(t)}{dt} + 2x(t). \tag{A.37}$$

The frequency response of this system is

$$H(j\omega) = \frac{j\omega + 2}{(j\omega)^2 + 4j\omega + 3}. \tag{A.38}$$

To determine the impulse response for this system, we expand $H(j\omega)$ into a sum of simpler terms whose inverse transforms can be obtained by inspection. Making the subsitution of v for $j\omega$, we obtain the function

---

[1] Here, we use the factorial notation $r!$ for the product $r(r-1)(r-2)\ldots 2 \cdot 1$. The quantity $0!$ is defined to be equal to 1.

$$G(v) = \frac{v+2}{v^2+4v+3} = \frac{v+2}{(v+1)(v+3)}. \tag{A.39}$$

The partial-fraction expansion for $G(v)$ is then

$$G(v) = \frac{A_{11}}{v+1} + \frac{A_{21}}{v+3}, \tag{A.40}$$

where

$$A_{11} = [(v+1)G(v)]\,|_{v=-1} = \frac{-1+2}{-1+3} = \frac{1}{2}, \tag{A.41}$$

$$A_{21} = [(v+3)G(v)]\,|_{v=-3} = \frac{-3+2}{-3+1} = \frac{1}{2}. \tag{A.42}$$

Thus,

$$H(j\omega) = \frac{\frac{1}{2}}{j\omega+1} + \frac{\frac{1}{2}}{j\omega+3}, \tag{A.43}$$

and the impulse response of the system, obtained by inverting eq. (A.43), is

$$h(t) = \frac{1}{2}e^{-t}u(t) + \frac{1}{2}e^{-3t}u(t). \tag{A.44}$$

The system described by eq. (A.37) can also be analyzed using the techniques of Laplace transform analysis, as developed in Chapter 9. The system function for this system is

$$H(s) = \frac{s+2}{s^2+4s+3}, \tag{A.45}$$

and if we substitute v for s, we obtain the same $G(v)$ given in eq. (A.39). Thus, the partial-fraction expansion proceeds exactly as in eqs. (A.40)–(A.42), with the result that

$$H(s) = \frac{\frac{1}{2}}{s+1} + \frac{\frac{1}{2}}{s+3}. \tag{A.46}$$

Inverting this transform, we again obtain the impulse response, as given in eq. (A.44).

Example A.2

We now illustrate the method of partial-fraction expansion when there are repeated factors in the denominator. In Example 4.26, we considered the response of the system described in eq. (A.37) when the input was

$$x(t) = e^{-t}u(t). \tag{A.47}$$

From eq. 4.81, the Fourier transform of the output of the system is

$$Y(j\omega) = \frac{j\omega+2}{(j\omega+1)^2(j\omega+3)}. \tag{A.48}$$

Substituting v for $j\omega$, we obtain the rational function

$$G(v) = \frac{v+2}{(v+1)^2(v+3)}. \tag{A.49}$$

The B_{ik} can be calculated in a manner similar to that used earlier:

$$\boxed{B_{ik} = \frac{1}{(\sigma_i - k)!}(-\rho_i)^{\sigma_i - k}\left[\frac{d^{\sigma_i - k}}{dv^{\sigma_i - k}}[(1 - \rho_i^{-1}v)^{\sigma_i}G(v)]\right]\Bigg|_{v=\rho_i}.} \tag{A.59}$$

As before, the validity of eq. (A.59) can be determined by multiplying both sides of eq. (A.58) by $(1 - \rho_i^{-1}v)^{\sigma_i}$, then differentiating repeatedly with respect to v, until B_{ik} is no longer multiplied by a power of $1 - \rho_i^{-1}v$, and finally, setting $v = \rho_i$.

Example A.3

Consider the causal LTI system in Example 5.19 characterized by the difference equation

$$y[n] - \frac{3}{4}y[n-1] + \frac{1}{8}y[n-2] = 2x[n]. \tag{A.60}$$

The frequency response of the system is

$$H(e^{j\omega}) = \frac{2}{1 - \frac{3}{4}e^{-j\omega} + \frac{1}{8}e^{-2j\omega}}. \tag{A.61}$$

For discrete-time transforms such as this, it is most convenient to substitute v for $e^{-j\omega}$. Making this substitution, we obtain the rational function

$$G(v) = \frac{2}{1 - \frac{3}{4}v + \frac{1}{8}v^2} = \frac{2}{(1 - \frac{1}{2}v)(1 - \frac{1}{4}v)}. \tag{A.62}$$

Using the partial-fraction expansion specified by eqs. (A.57)–(A.59), we obtain

$$G(v) = \frac{B_{11}}{1 - \frac{1}{2}v} + \frac{B_{21}}{1 - \frac{1}{4}v}, \tag{A.63}$$

$$B_{11} = \left[\left(1 - \frac{1}{2}v\right)G(v)\right]\Bigg|_{v=2} = \frac{2}{1 - \frac{1}{2}} = 4, \tag{A.64}$$

$$B_{21} = \left[\left(1 - \frac{1}{4}v\right)G(v)\right]\Bigg|_{v=4} = \frac{2}{1 - 2} = -2. \tag{A.65}$$

Thus,

$$H(e^{j\omega}) = \frac{4}{1 - \frac{1}{2}e^{-j\omega}} - \frac{2}{1 - \frac{1}{4}e^{-j\omega}}, \tag{A.66}$$

and taking the inverse transform of eq. (A.66), we obtain the unit impulse response:

$$h[n] = 4\left(\frac{1}{2}\right)^n u[n] - 2\left(\frac{1}{4}\right)^n u[n]. \tag{A.67}$$

In Section 10.7, we developed the tools of z-transform analysis for the examination of discrete-time LTI systems specified by linear constant-coefficient difference equations. Applying those techniques to this example, we find that the system function can be determined by inspection from eq. (A.60) and is

$$H(z) = \frac{2}{1 - \frac{3}{4}z^{-1} + \frac{1}{8}z^{-2}}. \tag{A.68}$$

Then, substituting v for z^{-1}, we obtain $G(v)$ as in eq. (A.62). Thus, using the partial-fraction expansion calculations in eqs. (A.63)–(A.65), we find that

$$H(z) = \frac{4}{1 - \frac{1}{2}z^{-1}} - \frac{2}{1 - \frac{1}{4}z^{-1}}, \tag{A.69}$$

which, when inverted, again yields the unit impulse response of eq. (A.67).

Example A.4

Suppose that the input to the system considered in Example A.3 is

$$x[n] = \left(\frac{1}{4}\right)^n u[n]. \tag{A.70}$$

Then from Example 5.20, the Fourier transform of the output is

$$Y(e^{j\omega}) = \frac{2}{(1 - \frac{1}{2}e^{-j\omega})(1 - \frac{1}{4}e^{-j\omega})^2}. \tag{A.71}$$

Substituting v for $e^{-j\omega}$ yields

$$G(v) = \frac{2}{(1 - \frac{1}{2}v)(1 - \frac{1}{4}v)^2}. \tag{A.72}$$

Thus, using eqs. (A.58) and (A.59), we obtain the partial-fraction expansion

$$G(v) = \frac{B_{11}}{1 - \frac{1}{4}v} + \frac{B_{12}}{(1 - \frac{1}{4}v)^2} + \frac{B_{21}}{1 - \frac{1}{2}v} \tag{A.73}$$

and find

$$B_{11} = (-4)\left[\frac{d}{dv}\left(1 - \frac{1}{4}v\right)^2 G(v)\right]\Bigg|_{v=4} = -4, \tag{A.74}$$

$$B_{12} = \left[\left(1 - \frac{1}{4}v\right)^2 G(v)\right]\Bigg|_{v=4} = -2, \tag{A.75}$$

$$B_{21} = \left[\left(1 - \frac{1}{2}v\right) G(v)\right]\Bigg|_{v=2} = 8. \tag{A.76}$$

Therefore,

$$Y(j\omega) = -\frac{4}{1 - \frac{1}{4}e^{-j\omega}} - \frac{2}{(1 - \frac{1}{4}e^{-j\omega})^2} + \frac{8}{1 - \frac{1}{2}e^{-j\omega}}, \tag{A.77}$$

which can be inverted by inspection as follows, using the Fourier transform pairs in Table 4.2:

$$y[n] = \left\{ -4\left(\frac{1}{4}\right)^n - 2(n+1)\left(\frac{1}{4}\right)^n + 8\left(\frac{1}{2}\right)^n \right\} u[n]. \tag{A.78}$$

Example A.5

Improper rational functions are often encountered in the analysis of discrete-time systems. To illustrate this, and also to show how they can be analyzed using the techniques developed here, consider the causal LTI system characterized by the difference equation

$$y[n] + \frac{5}{6}y[n-1] + \frac{1}{6}y[n-2] = x[n] + 3x[n-1] + \frac{11}{6}x[n-2] + \frac{1}{3}x[n-3].$$

The frequency response of this system is

$$H(e^{j\omega}) = \frac{1 + 3e^{-j\omega} + \frac{11}{6}e^{-j2\omega} + \frac{1}{3}e^{-j3\omega}}{1 + \frac{5}{6}e^{-j\omega} + \frac{1}{6}e^{-j2\omega}}. \tag{A.79}$$

Substituting v for $e^{-j\omega}$, we obtain

$$G(v) = \frac{1 + 3v + \frac{11}{6}v^2 + \frac{1}{3}v^3}{1 + \frac{5}{6}v + \frac{1}{6}v^2}. \tag{A.80}$$

This rational function can be written as the sum of a polynomial and a proper rational function:

$$G(v) = c_0 + c_1 v + \frac{b_1 v + b_0}{1 + \frac{5}{6}v + \frac{1}{6}v^2}. \tag{A.81}$$

Equating eqs. (A.80) and (A.81), and multiplying by $1 + \frac{5}{6}v + \frac{1}{6}v^2$, we obtain

$$\begin{aligned} 1 + 3v + \frac{11}{6}v^2 + \frac{1}{3}v^3 = {} & (c_0 + b_0) + \left(\frac{5}{6}c_0 + c_1 + b_1\right)v \\ & + \left(\frac{1}{6}c_0 + \frac{5}{6}c_1\right)v^2 + \frac{1}{6}c_1 v^3. \end{aligned} \tag{A.82}$$

Equating coefficients, we see that

$$\begin{aligned} \frac{1}{6}c_1 &= \frac{1}{3} \rightarrow c_1 = 2, \\ \frac{1}{6}c_0 + \frac{5}{6}c_1 &= \frac{11}{6} \rightarrow c_0 = 1, \\ \frac{5}{6}c_0 + c_1 + b_1 &= 3 \rightarrow b_1 = \frac{1}{6}, \\ c_0 + b_0 &= 1 \rightarrow b_0 = 0. \end{aligned} \tag{A.83}$$

Thus,

$$H(e^{j\omega}) = 1 + 2e^{-j\omega} + \frac{\frac{1}{6}e^{-j\omega}}{1 + \frac{5}{6}e^{-j\omega} + \frac{1}{6}e^{-j2\omega}}. \tag{A.84}$$

Also, we can use the method developed here to expand the proper rational function in eq. (A.81):

$$\frac{\frac{1}{6}v}{1+\frac{5}{6}v+\frac{1}{6}v^2} = \frac{\frac{1}{6}v}{(1+\frac{1}{3}v)(1+\frac{1}{2}v)} = \frac{B_{11}}{(1+\frac{1}{3}v)} + \frac{B_{21}}{(1+\frac{1}{2}v)}. \tag{A.85}$$

The coefficients are

$$B_{11} = \left.\left(\frac{\frac{1}{6}v}{1+\frac{1}{2}v}\right)\right|_{v=-3} = 1,$$

$$B_{21} = \left.\left(\frac{\frac{1}{6}v}{1+\frac{1}{3}v}\right)\right|_{v=-2} = -1.$$

Therefore, we find that

$$H(e^{j\omega}) = 1 + 2e^{-j\omega} + \frac{1}{1+\frac{1}{3}e^{-j\omega}} - \frac{1}{1+\frac{1}{2}e^{-j\omega}}, \tag{A.86}$$

and by inspection, we can determine the impulse response of this system:

$$h[n] = \delta[n] + 2\delta[n-1] + \left[\left(-\frac{1}{3}\right)^n - \left(-\frac{1}{2}\right)^n\right]u[n]. \tag{A.87}$$

ANSWERS

Chapter 1 Answers

1.1. $-0.5, -0.5, j, -j, j, 1+j, 1+j, 1-j, 1-j$

1.2. $5e^{j0}, 2e^{j\pi}, 3e^{-j\pi/2}, e^{-j\pi/3}, \sqrt{2}e^{j\pi/4}, 2e^{-j\pi/2}, \sqrt{2}e^{j\pi/4}, e^{j\pi/2}, e^{-j\pi/12}$.

1.3. **(a)** $P_\infty = 0, E_\infty = \frac{1}{4}$ **(b)** $P_\infty = 1, E_\infty = \infty$ **(c)** $P_\infty = \frac{1}{2}, E_\infty = \infty$
(d) $P_\infty = 0, E_\infty = \frac{4}{3}$ **(e)** $P_\infty = 1, E_\infty = \infty$ **(f)** $P_\infty = \frac{1}{2}, E_\infty = \infty$

1.4. **(a)** $n < 1$ and $n > 7$ **(b)** $n < -6$ and $n > 0$ **(c)** $n < -4$ and $n > 2$
(d) $n < -2$ and $n > 4$ **(e)** $n < -6$ and $n > 0$

1.5. **(a)** $t > -2$ **(b)** $t > -1$ **(c)** $t > -2$ **(d)** $t < 1$ **(e)** $t < 9$

1.6. **(a)** No **(b)** No **(c)** Yes

1.7. **(a)** $|n| > 3$ **(b)** all t **(c)** $|n| < 3, |n| \to \infty$ **(d)** $|t| \to \infty$

1.8. **(a)** $A = 2, a = 0, \omega = 0, \phi = \pi$ **(b)** $A = 1, a = 0, \omega = 3, \phi = 0$
(c) $A = 1, a = 1, \omega = 3, \phi = \frac{\pi}{2}$ **(d)** $A = 1, a = 2, \omega = 100, \phi = \frac{\pi}{2}$

1.9. **(a)** $T = \frac{\pi}{5}$ **(b)** Not periodic **(c)** $N = 2$
(d) $N = 10$ **(e)** Not periodic

1.10. π

1.11. 35

1.12. $M = -1, n_0 = -3$

1.13. 4

1.14. $A_1 = 3, t_1 = 0, A_2 = -3, t_2 = 1$

1.15. **(a)** $y[n] = 2x[n-2] + 5x[n-3] + 2x[n-4]$ **(b)** No

1.16. **(a)** No **(b)** 0 **(c)** No

1.17. **(a)** No; e.g., $y(-\pi) = x(0)$ **(b)** Yes

1.18. **(a)** Yes **(b)** Yes **(c)** $C \le (2n_0 + 1)B$

1.19. **(a)** Linear, not time invariant **(b)** Not linear, time invariant
(c) Linear, time invariant **(d)** Linear, not time invariant

1.20. **(a)** $\cos(3t)$ **(b)** $\cos(3t - 1)$

Chapter 2 Answers

2.1. **(a)** $y_1[n] = 2\delta[n+1] + 4\delta[n] + 2\delta[n-1] + 2\delta[n-2] - 2\delta[n-4]$
(b) $y_2[n] = y_1[n+2]$ **(c)** $y_3[n] = y_2[n]$

2.2. $A = n - 9, B = n + 3$

2.3. $2[1 - \frac{1}{2}^{n+1}]u[n]$

2.4.
$$y[n] = \begin{cases} n-6, & 7 \le n \le 11 \\ 6, & 12 \le n \le 18 \\ 24-n, & 19 \le n \le 23 \\ 0, & \text{otherwise} \end{cases}$$

2.5. $N = 4$

2.6. $y[n] = \begin{cases} \frac{3^n}{2}, & n < 0 \\ \frac{1}{2}, & n \geq 0 \end{cases}$

2.7. **(a)** $u[n-2] - u[n-6]$ **(b)** $u[n-4] - u[n-8]$ **(c)** No
(d) $y[n] = 2u[n] - \delta[n] - \delta[n-1]$

2.8. $y(t) = \begin{cases} t+3, & -2 < t \leq -1 \\ t+4, & -1 < t \leq 0 \\ 2-2t, & 0 < t \leq 1 \\ 0, & \text{elsewhere} \end{cases}$

2.9. $A = t - 5, B = t - 4$

2.10. **(a)** $y(t) = \begin{cases} t, & 0 \leq t \leq \alpha \\ \alpha, & \alpha \leq t \leq 1 \\ 1+\alpha-t, & 1 \leq t \leq 1+\alpha \\ 0, & \text{otherwise} \end{cases}$ **(b)** $\alpha = 1$

2.11. **(a)** $y(t) = \begin{cases} 0, & -\infty < t \leq 3 \\ \frac{1-e^{-3(t-3)}}{3}, & 3 < t \leq 5 \\ \frac{(1-e^{-6})e^{-3(t-5)}}{3}, & 5 < t \leq \infty \end{cases}$
(b) $g(t) = e^{-3(t-3)}u(t-3) - e^{-3(t-5)}u(t-5)$ **(c)** $g(t) = \frac{dy(t)}{dt}$

2.12. $A = \frac{1}{1-e^{-3}}$

2.13. **(a)** $A = \frac{1}{5}$ **(b)** $g[n] = \delta[n] - \frac{1}{5}\delta[n-1]$

2.14. $h_1(t), h_2(t)$

2.15. $h_2[n]$

2.16. **(a)** True **(b)** False **(c)** True **(d)** True

2.17. **(a)** $y(t) = \frac{1-j}{6}[e^{(-1+3j)t} - e^{-4t}]u(t)$
(b) $y(t) = \frac{1}{6}[e^{-t}(\cos 3t + \sin 3t) - e^{-4t}]u(t)$

2.18. $(1/4)^{n-1}u[n-1]$

2.19. **(a)** $\alpha = \frac{1}{4}, \beta = 1$ **(b)** $[2(\frac{1}{2})^n - (\frac{1}{4})^n]u[n]$

2.20. **(a)** 1 **(b)** 0 **(c)** 0

Chapter 3 Answers

3.1. $x(t) = 4\cos(\frac{\pi}{4}t) - 8\cos(\frac{3\pi}{4}t + \frac{\pi}{2})$

3.2. $x[n] = 1 + 2\sin(\frac{4\pi}{5}n + \frac{3\pi}{4}) + 4\sin(\frac{8\pi}{5}n + \frac{5\pi}{6})$

3.3. $\omega_0 = \frac{\pi}{3}, a_0 = 2, a_2 = a_{-2} = \frac{1}{2}, a_5 = a_{-5}^* = -2j$

3.4. $a_k = \begin{cases} 0, & k = 0 \\ e^{-jk\pi/2}\frac{3\sin(\frac{k\pi}{2})}{k\pi}, & k \neq 0 \end{cases}$

3.5. $\omega_2 = \omega_1, b_k = e^{-jk\omega_1}[a_{-k} + a_k]$

3.6. **(a)** $x_2(t), x_3(t)$ **(b)** $x_2(t)$

3.7. $a_k = \begin{cases} \frac{2}{T}, & k = 0 \\ \frac{b_k}{j\frac{2\pi}{T}k}, & k \neq 0 \end{cases}$

3.8. $x_1(t) = \sqrt{2}\sin(\pi t), x_2(t) = -\sqrt{2}\sin(\pi t)$

3.9. $a_0 = 3, a_1 = 1 - 2j, a_2 = -1, a_3 = 1 + 2j$

3.10. $a_0 = 0, a_{-1} = -j, a_{-2} = -2j, a_{-3} = -3j$

3.11. $A = 10, B = \frac{\pi}{5}, C = 0$

3.12. $c_k = 6$ for all k

3.13. $y(t) = 0$

3.14. $H(e^{j\pi/2}) = H^*(e^{j3\pi/2}) = 2e^{j\pi/4}, H(e^{j0}) = H(e^{j\pi}) = 0$

3.15. $|k| > 8$

3.16. **(a)** 0 **(b)** $\sin(\frac{3\pi}{8}n + \frac{\pi}{4})$ **(c)** 0

3.17. S_1 and S_3 are not LTI.

3.18. S_1 and S_2 are not LTI.

3.19. **(a)** $\frac{dy(t)}{dt} + y(t) = x(t)$ **(b)** $H(j\omega) = (\frac{1}{1+j\omega})$ **(c)** $y(t) = \frac{1}{\sqrt{2}}\cos(t - \frac{\pi}{4})$

3.20. **(a)** $\frac{d^2y(t)}{dt^2} + \frac{dy(t)}{dt} + y(t) = x(t)$ **(b)** $H(j\omega) = (\frac{1}{1+j\omega-\omega^2})$ **(c)** $-\cos t$

Chapter 4 Answers

4.1. **(a)** $\frac{e^{-j\omega}}{2+j\omega}$ **(b)** $\frac{4e^{-j\omega}}{4+\omega^2}$

4.2. **(a)** $2\cos\omega$ **(b)** $-2j\sin 2\omega$

4.3. **(a)** $\frac{\pi}{j}[e^{j\pi/4}\delta(\omega - 2\pi) - e^{-j\pi/4}\delta(\omega + 2\pi)]$
(b) $2\pi\delta(\omega) + \pi[e^{j\pi/8}\delta(\omega - 6\pi) + e^{-j\pi/8}\delta(\omega + 6\pi)]$

4.4. **(a)** $1 + \cos 4\pi t$ **(b)** $\frac{4j\sin^2 t}{\pi t}$

4.5. $x(t) = -\frac{2\sin(3(t-3/2))}{\pi(t-3/2)}, t = \frac{k\pi}{3} + \frac{3}{2}$ for nonzero integers k

4.6. **(a)** $X_1(j\omega) = 2X(-j\omega)\cos\omega$ **(b)** $X_2(j\omega) = \frac{1}{3}e^{-j2\omega}X(j\frac{\omega}{3})$
(c) $X_3(j\omega) = -\omega^2 e^{-j\omega}X(j\omega)$

4.7. **(a)** neither, neither **(b)** imaginary, odd **(c)** imaginary, neither
(d) real, even

4.8. **(a)** $\frac{2\sin(\omega/2)}{j\omega^2} + \pi\delta(\omega)$ **(b)** $\frac{2\sin(\omega/2)}{j\omega^2}$

4.9. **(a)** $\frac{\sin\omega}{j\omega^2} - \frac{e^{-j\omega}}{j\omega}$ **(b)** $\frac{\sin\omega}{\omega}$ **(c)** $\frac{\sin\omega}{j\omega^2} - \frac{\cos\omega}{j\omega}$

4.10. **(a)** $X(j\omega) = \begin{cases} j/2\pi, & -2 \leq \omega < 0 \\ -j/2\pi, & 0 \leq \omega < 2 \\ 0, & \text{otherwise} \end{cases}$ **(b)** $A = \frac{1}{2\pi^3}$

4.11. $A = \frac{1}{3}, B = 3$

4.12. **(a)** $-\frac{4j\omega}{(1+\omega^2)^2}$ **(b)** $-j2\pi\omega e^{-|\omega|}$

4.13. **(a)** No **(b)** Yes **(c)** Yes

4.14. $x(t) = \sqrt{12}[e^{-t} - e^{-2t}]u(t)$

4.15. $x(t) = 2te^{-|t|}u(t)$

4.16. **(a)** $g(t) = \pi \sum_{k=-\infty}^{\infty} \delta(t - \frac{k\pi}{4})$ **(b)** $X(j\omega) = \begin{cases} 4, & |\omega| \leq 1 \\ 0, & 1 < |\omega| \leq 4 \end{cases}$

4.17. **(a)** False **(b)** True

4.18. $h(t) = \begin{cases} \frac{5}{4}, & |t| < 1 \\ -\frac{|t|}{4} + \frac{3}{2}, & 1 \leq |t| \leq 5 \\ -\frac{|t|}{8} + \frac{7}{8}, & 5 < |t| < 7 \\ 0, & \text{otherwise} \end{cases}$

4.19. $x(t) = e^{-4t}u(t)$

4.20. $h(t) = \frac{2}{\sqrt{3}}e^{-t/2}\sin(\frac{\sqrt{3}}{2}t)u(t)$

Chapter 5 Answers

5.1. **(a)** $\frac{e^{-j\omega}}{1-\frac{1}{2}e^{-j\omega}}$ **(b)** $\frac{0.75e^{-j\omega}}{1.25-\cos\omega}$

5.2. **(a)** $2\cos\omega$ **(b)** $2j\sin(2\omega)$

5.3. **(a)** $\frac{\pi}{j}\{e^{j\pi/4}\delta(\omega - \frac{\pi}{3}) - e^{-j\pi/4}\delta(\omega + \frac{\pi}{3})\}$
(b) $4\pi\delta(\omega) + \pi\{e^{j\pi/8}\delta(\omega - \frac{\pi}{6}) + e^{-j\pi/8}\delta(\omega + \frac{\pi}{6})\}$

5.4. **(a)** $x_1[n] = 1 + \cos(\frac{\pi}{2}n)$ **(b)** $-4\frac{\sin^2(\frac{\pi}{2}n)}{\pi n}$

5.5. $x[n] = \frac{\sin[\frac{\pi}{4}(n-\frac{3}{2})]}{\pi(n-\frac{3}{2})}$, and $x[n] = 0$ for $n = \pm\infty$

5.6. **(a)** $X_1(e^{j\omega}) = (2\cos\omega)X(e^{-j\omega})$ **(b)** $X_2(e^{j\omega}) = \Re e\{X(e^{j\omega})\}$
(c) $X_3(e^{j\omega}) = -\frac{d^2}{d\omega^2}X(e^{j\omega}) - 2j\frac{d}{d\omega}X(e^{j\omega}) + X(e^{j\omega})$

5.7. **(a)** imaginary, neither **(b)** real, odd **(c)** real, neither

5.8. $x[n] = \begin{cases} 1, & n \leq -2 \\ n + 3, & -1 \leq n \leq 1 \\ 4, & n \geq 2 \end{cases}$

5.9. $x[n] = -\delta[n + 2] + \delta[n + 1] + \delta[n]$

5.10. $A = 2$

5.11. $\alpha = \pi$

5.12. $\frac{\pi}{2} \leq |\omega_c| \leq \pi$

5.13. $h_2[n] = -2(\frac{1}{4})^n u[n]$

5.14. $h[n] = \frac{16}{17}\delta[n] - \frac{1}{17}\delta[n - 2]$

5.15. $\omega_c = 3\pi/4$

5.16. **(a)** $\alpha = \frac{1}{4}$ **(b)** $N = 4$ **(c)** No

5.17. $b_k = \frac{1}{2}(-1)^k$

5.18. $a_k = \frac{1}{3}(\frac{1}{2})^{|k|}$

5.19. **(a)** $H(e^{j\omega}) = \frac{1}{(1-\frac{1}{2}e^{-j\omega})(1+\frac{1}{3}e^{-j\omega})}$

(b) $h[n] = \frac{3}{5}(\frac{1}{2})^n u[n] + \frac{2}{5}(-\frac{1}{3})^n u[n]$

5.20. **(a)** $H(e^{j\omega}) = \frac{\frac{4}{5}e^{-j\omega}}{1-\frac{4}{5}e^{-j\omega}}$

(b) $y[n] - \frac{4}{5}y[n-1] = \frac{4}{5}x[n-1]$

Chapter 6 Answers

6.1. **(a)** $A = |H(j\omega_0)|$ **(b)** $t_0 = -\frac{\sphericalangle H(j\omega_0)}{\omega_0}$

6.2. $\sphericalangle H(e^{j\omega_0}) = -n_0(\omega_0) + 2\pi k$ for some integer k.

6.3. **(a)** $A = 1$ **(b)** $\tau(\omega) > 0$ for $\omega > 0$

6.4. **(a)** $2\cos(\frac{\pi}{2}n - \pi)$ **(b)** $2\sin(\frac{7\pi}{2}n - \frac{3\pi}{4})$

6.5. **(a)** $g(t) = 2\cos(2\omega_c t)$ **(b)** more concentrated

6.6. **(a)** $g[n] = (-1)^n$ **(b)** more concentrated

6.7. **(a)** 1,000 Hz and 3,000 Hz **(b)** 800 Hz and 3,200 Hz

6.8. $\pi - \omega_p \leq \omega \leq \pi$

6.9. Final value $= 2/5$, $t_0 = 2/5$ sec

6.10. **(a)** $$20\log_{10}|H(j\omega)| \approx \begin{cases} -20, & \omega \ll 0.1 \\ 20\log_{10}(\omega), & 0.1 \ll \omega \ll 40 \\ 32, & \omega \gg 40 \end{cases}$$

(b) $$20\log_{10}|H(j\omega)| \approx \begin{cases} 20, & \omega \ll 0.2 \\ -20\log_{10}(\omega) + 6, & 0.2 \ll \omega \ll 50 \\ -28, & \omega \gg 50 \end{cases}$$

6.11. **(a)** $$20\log_{10}|H(j\omega)| \approx \begin{cases} 20, & \omega \ll 0.5 \\ -20\log_{10}(\omega) + 14, & 0.5 \ll \omega \ll 50 \\ -40\log_{10}(\omega) + 48, & \omega \gg 50 \end{cases}$$

(b) $$20\log_{10}|H(j\omega)| \approx \begin{cases} 0, & \omega \ll 1 \\ -40\log_{10}\omega, & 1 \ll \omega \ll 50 \\ -20\log_{10}\omega - 34, & \omega \gg 50 \end{cases}$$

6.12. $H_2(j\omega) = \frac{0.01(j\omega+40)}{(j\omega+1)(j\omega+8)}$

6.13. **(a)** not unique **(b)** unique

6.14. $H_I(j\omega) = 0.2 \times 10^{-4}\frac{(j\omega+50)(j\omega+10)}{(j\omega+0.2)^2}$

6.15. **(a)** critically damped **(b)** underdamped
(c) overdamped **(d)** underdamped

6.16. $y[n] + \frac{1}{2}y[n-1] = \frac{3}{2}x[n]$

6.17. **(a)** oscillatory **(b)** nonoscillatory

6.18. No

6.19. $R \geq 2\sqrt{\frac{L}{C}}$

6.20. $\tau(\omega) = 2$

Chapter 7 Answers

7.1. $|\omega| > 5{,}000\pi$

7.2. (a) and (c)

7.3. **(a)** $8{,}000\pi$ **(b)** $8{,}000\pi$ **(c)** $16{,}000\pi$

7.4. **(a)** ω_0 **(b)** ω_0 **(c)** $2\omega_0$ **(d)** $3\omega_0$

7.5. $|H(j\omega)| = \begin{cases} T, & |\omega| \leq \omega_c \\ 0, & \text{otherwise} \end{cases}$, where $\frac{\omega_0}{2} < \omega_c < \frac{2\pi}{T} - \frac{\omega_0}{2}$, $\sphericalangle H(j\omega) = 0$

7.6. $T_{\max} = \frac{\pi}{\omega_1 + \omega_2}$

7.7. $H(j\omega) = \frac{2\sin(\omega T/2)}{\omega T} \times e^{j(\omega T/2)}$

7.8. **(a)** Yes

(b) $g(t) = \sum_{k=-4}^{4} a_k e^{jk\pi t}$, where $a_k = \begin{cases} 0, & k = 0 \\ -j(\frac{1}{2})^{k+1}, & 1 \leq k \leq 4 \\ j(\frac{1}{2})^{-k+1}, & -4 \leq k \leq -1 \end{cases}$

7.9. $\omega_0 = 50\pi$

7.10. **(a)** False **(b)** True **(c)** True

7.11. **(a)** $X_c(j\omega)$ is real
(b) $\text{Max}\{X_c(j\omega)\} = 0.5 \times 10^{-3}$
(c) $X_c(j\omega) = 0$ for $|\omega| \geq 1{,}500\pi$
(d) $X_c(j\omega) = X_c(j(\omega - 2{,}000\pi))$ for $0 \leq \omega \leq 2{,}000\pi$

7.12. $|\omega| \geq 750\pi$

7.13. $h[n] = \delta[n-2]$

7.14. $h[n] = -\frac{\sin[\pi(n-\frac{1}{2})]}{T\pi(n-\frac{1}{2})^2}$

7.15. $N = 2$

7.16. $x[n] = 4(\frac{\sin(\pi n/2)}{\pi n})^2$

7.17. Ideal lowpass filter with cutoff frequency $\pi/2$ and passband gain of unity

7.18. Ideal lowpass filter with cutoff frequency $\pi/4$ and passband gain of 2.

7.19. **(a)** $y[n] = \frac{\sin(5\omega_1 n/3)}{5\pi n}$ **(b)** $y[n] = \frac{1}{5}\delta[n]$

7.20. **(a)** Yes **(b)** No

Chapter 8 Answers

8.1. $m(t) = \frac{1}{2}e^{-j\omega_c t}$

8.2. **(a)** No constraint necessary **(b)** $|\omega_c| > 1{,}000\pi$

8.3. $y(t) = 0$

8.4. $y(t) = \sin 200\pi t$

8.5. $m = \frac{3}{2\pi}$

8.6. $A = 4$

8.7. $\omega_0 = 2\omega_c, A = 2$

8.8. **(a)** Yes **(b)** Yes, $x(t) = \{y(t)\sin\omega_c t\} * \frac{2\sin\omega_c t}{\pi t}$

8.9. **(a)** $|\omega| > 2\omega_c$ **(b)** $\omega_0 = \omega_c, A = 2$

8.10. **(a)** $X(j\omega) = 0$ for $|\omega| \geq 1{,}000\pi$ **(b)** $\omega_c = 1{,}000\pi, A = 4$

8.11. **(a)** $\frac{\omega_c}{2} \leq |\omega| \leq \frac{3\omega_c}{2}$, Gain $= 1$ **(b)** $A = 2|a_1|, \phi = \sphericalangle a_1$

8.12. $\Delta = 0.5 \times 10^{-4}$

8.13. **(a)** $p(0) = \frac{1}{T_1}$ **(b)** $p(kT_1) = 0$

8.14. $Y(j\omega) = \pi\delta(\omega - \omega_c) - \frac{m\pi}{2j}\delta(\omega - \omega_c - \omega_m) - \frac{m\pi}{2j}\delta(\omega - \omega_c + \omega_m)$

8.15. $\omega_0 = 0$ and $\omega_0 = \pi$

8.16. $0 \leq \omega \leq \frac{3\pi}{8}$ and $\frac{5\pi}{8} \leq \omega \leq \pi$

8.17. $0 \leq |\omega| \leq \frac{\pi}{2}$

8.18. $H(e^{j\omega}) = \begin{cases} j, & 0 < \omega \leq \frac{\pi}{4} \\ -j, & -\frac{\pi}{4} \leq \omega < 0 \end{cases}$

8.19. $N = 20$

8.20. $p[n] = \sum_{k=-\infty}^{\infty} \delta[n - 2k]$

Chapter 9 Answers

9.1. **(a)** $\sigma > -5$ **(b)** $\sigma < -5$ **(c)** $-\infty \leq \sigma \leq \infty$
(d) no value of σ **(e)** $|\sigma| < 5$ **(f)** $\sigma < 5$

9.2. **(a)** $\frac{e^{-(s+5)}}{s+5}, \mathcal{R}e\{s\} > -5$ **(b)** $A = -1, t_0 = -1, \mathcal{R}e\{s\} < -5$

9.3. $\mathcal{R}e\{\beta\} = 3, \mathcal{I}m\{\beta\}$ arbitrary

9.4. $1 + 2j, 1 - 2j, \mathcal{R}e\{s\} < 1$

9.5. **(a)** 1,1 **(b)** 0,1 **(c)** 1,0

9.6. **(a)** no **(b)** yes **(c)** no **(d)** yes

9.7. 4

9.8. two sided

9.9. $x(t) = 4e^{-4t}u(t) - 2e^{-3t}u(t)$

9.10. **(a)** lowpass **(b)** bandpass **(c)** highpass

9.11. $|X(j\omega)| = 1$

9.12. **(a)** not consistent **(b)** consistent **(c)** consistent

9.13. $\alpha = -1, \beta = \frac{1}{2}$

9.14. $X(s) = 1/[4(s^2 - \frac{s}{\sqrt{2}} + \frac{1}{4})(s^2 + \frac{s}{\sqrt{2}} + \frac{1}{4})], -\frac{\sqrt{2}}{4} < \mathcal{R}e\{s\} < \frac{\sqrt{2}}{4}$

9.15. $X(s) = \frac{s}{s^2+4}; \mathcal{R}e\{s\} > 0, \qquad Y(s) = \frac{2}{s^2+4}; \mathcal{R}e\{s\} > 0$

9.16. **(a)** 2 **(b)** $\alpha > 0$

9.17. $\frac{d^2y(t)}{dt^2} + 10\frac{dy(t)}{dt} + 16y(t) = 12x(t) + 3\frac{dx(t)}{dt}$

9.18. **(a)** $H(s) = \frac{1}{s^2+s+1}, \mathcal{R}e\{s\} > -\frac{1}{2}$
(b) Lowpass
(c) $H(s) = \frac{1}{s^2+10^{-3}s+1}, \mathcal{R}e\{s\} > -0.0005$
(d) Bandpass

9.19. **(a)** $\frac{1}{s+2}, \mathcal{R}e\{s\} > -2$
(b) $1 + \frac{e^{-6}}{s+2}, \mathcal{R}e\{s\} > -2$
(c) $\frac{1}{s+4} + \frac{1}{s+2}, \mathcal{R}e\{s\} > -2$

9.20. **(a)** $e^{-t}u(t) - e^{-2t}u(t)$
(b) $e^{-t}u(t)$
(c) $2e^{-t}u(t) - e^{-2t}u(t)$

Chapter 10 Answers

10.1. **(a)** $|z| > \frac{1}{2}$ **(b)** $|z| < \frac{1}{2}$ **(c)** $|z| > 1$ **(d)** $\frac{1}{2} < |z| < 2$

10.2. $X(z) = \frac{1}{125}\frac{z^{-3}}{1-\frac{1}{5}z^{-1}}; |z| > \frac{1}{5}$

10.3. $|\alpha| = 2$, n_0 arbitrary

10.4. poles at $z = \frac{1}{3}e^{\pm j\pi/4}$, ROC: $|z| < \frac{1}{3}$

10.5. **(a)** 1,1 **(b)** 2,0 **(c)** 1,2

10.6. **(a)** No **(b)** No **(c)** Yes **(d)** Yes

10.7. 3

10.8. two sided

10.9. $x[n] = \frac{2}{9}u[n] + \frac{7}{9}(-2)^n u[n]$

10.10. **(a)** $x[0] = 1, x[1] = \frac{2}{3}, x[2] = -\frac{2}{9}$ **(b)** $x[0] = 3, x[-1] = -6, x[-2] = 18$

10.11. $x[n] = \begin{cases} (\frac{1}{2})^n, & 0 \leq n \leq 9 \\ 0, & \text{otherwise} \end{cases}$

10.12. **(a)** highpass **(b)** lowpass **(c)** bandpass

10.13. **(a)** $G(z) = 1 - z^{-6}; |z| > 0$ **(b)** $X(z) = \frac{1-z^{-6}}{1-z^{-1}}; |z| > 0$

10.14. **(a)** $n_0 = 2$ **(b)** $G(z) = (\frac{z^{-1}-z^{-7}}{1-z^{-1}})^2$

10.15. $(\frac{1}{3})^n u[n]$ and $(-\frac{1}{3})^n u[n]$

10.16. **(a)** Not causal **(b)** Causal **(c)** Not causal

10.17. **(a)** Yes **(b)** Yes

10.18. **(a)** $y[n] - \frac{2}{3}y[n-1] + \frac{1}{9}y[n-2] = x[n] - 6x[n-1] + 8x[n-2]$
(b) Yes

10.19. **(a)** $\chi_1(z) = \frac{1}{1-\frac{1}{4}z^{-1}}, |z| > \frac{1}{4}$
(b) $\chi_2(z) = 2$, All z
(c) $\chi_3(z) = \frac{1}{1-\frac{1}{2}z^{-1}}, |z| > \frac{1}{2}$

10.20. **(a)** $-(-\frac{1}{2})^n u[n]$
(b) $\frac{1}{3}(-\frac{1}{2})^n u[n] + \frac{1}{6}(\frac{1}{4})^n u[n]$
(c) $-\frac{2}{3}(-\frac{1}{2})^n u[n] + \frac{1}{6}(\frac{1}{4})^n u[n]$

Chapter 11 Answers

11.1. $H_0(z) + \frac{H_1(z)}{1+G(z)H_1(z)}$

11.2. $\frac{H_1(s)H_2(s)}{1+H_1(s)G_1(s)+H_1(s)H_2(s)G_2(s)}$

11.3. $b < -1$

11.4. $G(s) = \frac{1}{s}$

11.5. $-\frac{5}{2} < b < \frac{3}{2}$

11.6. FIR

11.7. $K > -6$

11.8. $-3 < k < 0$

11.9. No, root locus stays on real axis

11.10. Double pole at $s = -1$, double zero at $s = 1$

11.11. $0 < k < \frac{5}{4}$

11.12. Pole and zero positions alternate on the real axis

11.13. Unstable for all K

11.14. **(a)** 0 **(b)** 1

11.15. $K > -1$

11.16. $K > -1$

11.17. $-1 < K < 4$

11.18. $-1 < K < 1$

11.19. Unstable

11.20. Gain margin is infinite, phase margin is $2\tan^{-1}\sqrt{2}$

INDEX

Absolutely summable impulse response, 113
Absolutely integrable impulse response, 114
Accumulation property
  discrete-time Fourier series, 221
  discrete-time Fourier transform, 375–76
  unilateral z transform, 793
Accumulator, 44
Acoustic feedback, 830–32, 855
Adders in block diagrams, 125, 126
Additivity property, 53
Aliasing, 527–34
All-pass systems, 430, 498, 681–82
AM. *See* Amplitude modulation (AM)
Amplifier
  chopper, 652
  operational, 821, 896–97
Amplitude modulation (AM), 236–37, 322, 324, 583
  pulse-train carrier, 601–4, 605
  sinusoidal, 583–87
    complex exponential carrier, 583–85
    demodulation for, 587–94
    discrete-time, 619–23
    frequency-division multiplexing (FDM) using, 594–97
    single-sideband, 597–601
    sinusoidal carrier, 585–87
Amplitude-scaling factor, 483
Analog-to-digital (A-to-D) converter, 535
Analysis equation
  continuous-time Fourier series, 191
  continuous-time Fourier transform, 288
  discrete-time Fourier series, 213
  discrete-time Fourier transform, 361, 390
Angle criterion, 836–40
Angle modulation, 611–13
Angle (phase) of complex number, 71
Anticausality, 695
Aperiodic convolution, 222
Aperiodic signal, 12, 180
  continuous-time Fourier transform for, 285–89
  discrete-time Fourier transform for, 359–62
Associative property of LTI systems, 107–8
Audio systems
  feedback in, 830–32, 855
  frequency-shaping filters in, 232
Autocorrelation functions, 65, 168, 170–72, 738
Automobile suspension system, analysis of, 473–76
Average, weighted, 245
Averaging system, noncausal, 47

Band-limited input signals, 541
Band-limited interpolation, 523–24
Bandpass filters, 237–38, 326
Bandpass-sampling techniques, 564–65

Bandpass signal, 564–65
Bandwidth of an LTI system, 352–53
Bartlett (triangular) window, 421
Bernoulli, D., 178
Bilateral Laplace transform. *See* Laplace transform
Bilinear transformation, 814–15
Bit, 610
Block diagram(s), 42
  cascade-form, 712, 713, 787–89
  causal LTI systems, 708–13, 784–89
  direct-form, 712, 713, 787–89
  first-order systems described by differential and difference equations, 124–27
  parallel-form, 712, 713, 787–89
Bode plots, 436–39
  automobile suspension system, 475
  rational frequency responses, 456–60
Break frequency, 450
Butterworth filters, 446–47, 505, 703–6

Capacitor, 44
Carrier frequency, 584
Carrier signal, 583
Cartesian (rectangular) form for complex number, 71
Cascade-form block diagrams, 712, 713, 787–89
Cascade (series) interconnection, 42
Causal LTI systems, 46–48, 112–13, 116–27
  block diagram representations for, 708–13, 784–89
    first-order systems, 124–27
  Laplace transform for, 693–95, 697
  z-transform for, 776–77
Channel equalization, 609–10
“Chirp” transform algorithm, 651
Chopper amplifier, 652
Circle, unit, 743
Circuit, quality of, 456
Closed-loop system, 818
Closed-loop poles, 834–36
Closed-loop system function, 820
Coefficient multiplier, 125, 126
Coefficients, Fourier series. *See* Fourier series coefficients
Communications systems, 582–653
  amplitude modulation with pulse-train carrier, 601–4, 605
  discrete-time modulation, 619–23
  pulse-amplitude modulation, 604–10
    digital, 610
    intersymbol interference in, 607–10
  sinusoidal amplitude modulation, 583–87
    with complex exponential carrier, 583–85
    demodulation for, 587–94
    discrete-time, 619–23
    frequency-division multiplexing (FDM) using, 594–97
    single-sideband, 597–601
    sinusoidal carrier, 585–87
  sinusoidal frequency modulation, 583, 611–19
    narrowband, 613–15
    periodic square-wave modulating signal, 617–19
    wideband, 615–17
Commutative property of LTI systems, 104
Compensation for nonideal elements, 821–22
Complex conjugate, 72
Complex exponential(s)
  discrete-time, periodicity properties of, 25–30
  general, 20–21
  harmonically related, 19
  linear combinations of harmonically related. *See* Fourier series
  LTI system response to, 182–86, 226–31
  periodic, 186
  sinusoidal amplitude modulation and, 583–85
Complex exponential signals, general, 24–25

Complex numbers, 71
Conditionally stable systems, 892
Conjugate symmetry, 204–5, 206, 221, 303–6, 375
Conjugation property
  continuous-time Fourier series, 204–5, 206
  continuous-time Fourier transform, 303–6
  discrete-time Fourier series, 221
  discrete-time Fourier transform, 375
  Laplace transform, 687
    unilateral, 717
  z-transform, 770
    unilateral, 793
Constants, time, 448
  dominant, 500–501
Continuous-time signals
  energy and power of, 5–7
  examples and mathematical representation of, 1–5
  sampling of, 4
Continuous-time systems, 448–60
  Bode plots for rational frequency responses and, 456–60
  examples of, 39–41
  first-order, 448–51
  interconnections of, 41–43
  second-order, 451–56
Continuous-to-discrete-time conversion, 535
Convergence. *See also* Region of convergence
  continuous-time Fourier series, 195–201
  continuous-time Fourier transform, 289–90
  discrete-time Fourier series, 219–21
  discrete-time Fourier transform, 366–67
Convolution
  aperiodic, 222
  associative property of, 107–8
  commutative property of, 104
  defining continuous-time unit impulse function through, 131–32
  distributive property of, 104–6
  operation of, 79, 85
  periodic, 222, 389–90
Convolution integral, 90–102
  evaluating, 97–102
Convolution property
  continuous-time Fourier series, 206
  continuous-time Fourier transform, 314–22
  discrete-time Fourier series, 221, 222
  discrete-time Fourier transform, 382–88
  Laplace transform, 687–88
    unilateral, 717–18
  z-transform, 770–72
    unilateral, 793–94
Convolution sum, 75–90, 384–86
  evaluating, 81–84
Correlation function, 65
Critically damped systems, 453, 467
Cross-correlation functions, 65, 168, 170
Cutoff frequencies, 237

Damped sinusoids, 21
Damping ratio, 452–53
dc offset, 207
dc sequence, 224
Deadbeat feedback, 907
Decibels (dB), 232, 233, 437
Decimation, 549–55
Degenerative (negative) feedback, 822, 831–32
Delay, 44
  group, 430–36
  half-sample, 543–45
  unit, 125
Delay time, 406
Demodulation, 585
  defined, 583
  sinusoidal amplitude modulation, 587–94
    asynchronous, 590–94
    discrete-time, 622
    synchronous, 587–90

Difference equations
  discrete-time filters described by, 244–49
    first-order recursive, 244–45
    nonrecursive, 245–49
  linear constant-coefficient. *See* Linear constant-coefficient difference equations
Differencing property
  discrete-time Fourier series, 222–23
  discrete-time Fourier transform, 375–76
  z-transform, 775
    unilateral, 793
Differential equations
  continuous-time filters described by, 239–44
    RC highpass filter, 241–44
    RC lowpass filter, 239–41
  linear constant-coefficient. *See* Linear constant-coefficient differential equations
Differentiating filters
  continuous-time, 232–34, 235
  discrete-time, 541–43
Differentiation
  s domain, 688–90, 717
  time-domain, 688, 717
  z domain, 772, 793
Differentiation property
  continuous-time Fourier series, 206
  continuous-time Fourier transform, 306–8
  discrete-time Fourier transform, 380
Differentiator
  block diagram representation of, 126
  digital, 541–43
Digital-to-analog (D-to-A) converter, 535
Direct Form I realization, 161–63
Direct Form II realization, 161–63
Direct-form block diagrams, 712, 713, 787–89
Dirichlet, P.L., 180
Dirichlet conditions, 197–200, 290, 316
Discontinuities, 198–200
Discrete Fourier transform (DFT) for finite-duration signals, 417–18
Discrete-time Fourier series pair, 213
Discrete-time LTI filters, 234–36
Discrete-time modulation, 619–23
Discrete-time nonrecursive filters, 476–82
Discrete-time signals, 211
  energy and power of, 5–7
  examples and mathematical representation of, 1–5
  sampling of, 545–55
    decimation and interpolation, 549–55
    impulse train, 545–49
Discrete-time systems, 461–72
  first-order, 461–65
  second-order, 465–72
Discrete-time to continuous-time conversion, 535
Dispersion, 433
Distortion
  magnitude and phase, 428
  quadrature, 636
Distribution theory, 36, 136
Distributive property of LTI systems, 104–6
Dominant time constant, 500–501
Double-sideband modulation (DSB), 598
Downsampling, 551
Duality
  continuous-time Fourier transform, 295, 309–11, 322
discrete-time Fourier transform, 390–96
  between discrete-time Fourier transform and continuous-time Fourier series, 395–96

Echo, 737
Eigenfunctions, 183–84, 272
Eigenvalue, 183, 272
Elliptic filters, 446–47
Encirclement property, 847–50
Energy-density spectrum, 312, 349, 381
Energy of signals, 5–7
Envelope detector, 591, 592
Envelope function, 285
Equalization, 649
  channel, 609–10

Equalizer, zero-force, 649
Equalizer circuits, 232, 233
Euler, L., 178
Euler's relation, 71
Even signals, 13–14
  continuous-time Fourier series, 206
  discrete-time Fourier series, 221
Exponentials. *See* Complex exponential(s)
Exponential signals, 14–30
  continuous-time complex, 15–21
  discrete-time complex, 21–30
  real, 22, 23
Extension problems, 137

Fast Fourier transform (FFT), 182, 418
Feedback. *See also* Linear feedback systems
  applications of, 820–32
    angular position of telescope, 816–18
    audio, 830–32, 855
    compensation for nonideal elements, 821–22
    inverse system design, 820–21
    inverted pendulum, 818–19
    population dynamics, 824–26
    sampled-data systems, 826–28
    stabilization of unstable systems, 823–26
    tracking systems, 828–30
  closed-loop poles, 836–38
  deadbeat, 907
  degenerative (negative), 822, 831–32
  destabilization caused by, 830–32
  positive (regenerative), 831–32
  proportional, 823
  proportional-plus-derivative, 824
  Type *l*, 906
Feedback interconnection, 43
Feedback path, system function of, 820
Filter(s), 231–50
  Butterworth, 446–47, 505, 703–6
  continuous-time, described by differential equations, 239–44
    RC highpass filter, 241–44
    RC lowpass filter, 239–41
  continuous-time differentiating, 232–34, 235, 541
  discrete-time, described by difference equations, 244–49
    first-order recursive, 244–45
    nonrecursive, 245–49
  elliptic, 446–47
  finite impulse response (FIR), 122
    discrete-time, 476–82
    linear-phase, causal, symmetric, 579–80
  frequency-selective, 231, 236–39, 318, 423
    bandpass, 237–38, 326
    highpass, 237–38, 241–44, 374–75, 387
    ideal, 237–39, 439–44
    lowpass, 237, 318, 321–22, 374–75, 384, 440, 442, 443
    time-domain properties of ideal, 439–44
    with variable center frequency, 325–27
  frequency-shaping, 231, 232–36
    differentiating filters, 232–34, 235
    discrete-time LTI filters, 234–36
  infinite impulse response (IIR), 123, 476
  matched, 170–72, 275
  moving-average, 476–82
  nonideal, 444–47
Filtering, defined, 231
Final-value theorem, 690–91
Finite-duration signals, discrete Fourier transform for, 417–18
Finite impulse response (FIR) filters, 122
  discrete-time, 245–49, 476–82
  linear-phase, causal, symmetric, 579–80
Finite sum formula, 73
First difference, 111. *See also* Differencing property
First harmonic components, 187
First-order continuous-time systems, 448–51
First-order discrete-time systems, 461–65

First-order recursive discrete-time filters, 244–45
Forced response, 118
Forward path, system function of, 820
Fourier, Jean Baptiste Joseph, 179–81
Fourier series, 177–283. *See also* Filter(s)
  continuous-time, 186–211
    analysis equation of, 191
    conjugation and conjugate symmetry of, 204–5
    convergence of, 195–201
    determination of, 190–95
    Dirichlet conditions, 197–200
    duality between discrete-time Fourier transform and, 395–96
    examples, 205–11
    Gibbs phenomenon, 200–201, 219
    linear combinations of harmonically related complex exponentials, 186–90
    linearity property of, 202
    multiplication property of, 204
    Parseval's relation for, 205, 206
    square wave, 193–95, 200, 201, 209, 218–19
    synthesis equation of, 191
    table of properties, 206
    time reversal property of, 203
    time scaling property of, 204
    time shifting property of, 202–3
  discrete-time, 211–26
    analysis equation of, 213
    convergence of, 219–21
    determination of, 212–21
    first-difference property of, 222–23
    linear combinations of harmonically related complex exponentials, 211–12
    multiplication property of, 222
    Parseval's relation for, 223
    square wave, 219–20, 224
    synthesis equation of, 213
    table of properties, 221
  historical perspective, 178–82
  LTI systems and, 226–31
Fourier series coefficients
  continuous-time, 191, 286
  convolution property of, 222
  discrete-time, 212, 213
  real and imaginary parts of, 216, 217
Fourier series pair, discrete-time, 213
Fourier transform, 180
  fast (FFT), 182, 418
  geometric evaluation from pole-zero plot, 674–82, 763–67
    all-pass systems, 681–82
    first-order systems, 676–77, 763–65
    second-order systems, 677–81, 765–67
  inverse, 284, 288
  magnitude-phase representation of, 423–27
Fourier transform, continuous-time, 284–357. *See also* Laplace transform
  analysis equations, 300
  for aperiodic signal, 285–89
  convergence of, 289–90
  Dirichlet conditions, 290, 316
  even signals, 304
  examples of, 290–96
  Gibbs phenomenon, 294
  imaginary part of, 304
  impulse train, 299–300
  inverse, 284, 288
  odd signals, 304
  periodic signals, 296–300
  properties of, 300–330
    conjugation and conjugate symmetry, 303–6
    convolution, 314–22
    differentiation and integration, 306–8
    duality, 295, 309–11, 322
    linearity, 301
    multiplication, 322–27
    Parseval's relation, 312–14
    tables of, 328–30
    time and frequency scaling, 308–9
    time shifting, 301–3
  real part of, 304

- rectangular pulse signal, 293–94
- sinc functions, 295
- symmetric periodic square wave, 297–98
- synthesis equation for, 288, 300, 314
- two-dimensional, 356–57
- of unit impulse, 292

Fourier transform, discrete-time, 358–422
- analysis equation, 361, 390
- convergence issues associated with, 366–67
- development of, 359–62
- examples of, 362–66
- finite-duration signals, 417–18
- impulse train, 371–72, 376
- periodic signals, 367–72
- properties of, 372–90
  - conjugation and conjugation symmetry, 375
  - convolution, 382–88
  - differencing and accumulation, 375–76
  - differentiation in frequency, 380
  - duality, 390–96
  - linearity, 373
  - multiplication, 388–90
  - Parseval's relation, 380–82
  - periodicity, 373
  - table of, 390, 391
  - time expansion, 377–80
  - time reversal, 376–77
  - time shifting and frequency shifting, 373–75
- rectangular pulse, 365–66
- synthesis equation, 361, 390
- unit impulse, 367
- z-transform and, 743

Fourier transform pairs, 288
- continuous-time, 329
- discrete-time, 361, 392

Frequency(ies)
- carrier, 584
- cutoff, 237
- differentiation in, 380
- fundamental, 17–18
- instantaneous, 613
- Nyquist, 519
- passband, 237
- sampling, 516
- stopband, 237

Frequency-division multiplexing (FDM), 594–97

Frequency-domain characterization. *See* Time-domain and frequency-domain characterization

Frequency response, 227–28
- continuous-time delay, 543
- continuous-time ideal band-limited differentiator, 541
- discrete-time delay, 543
- discrete-time filter, 542
- first-order system, 677
- highpass filter, 374–75
- lowpass filter, 374–75
  - discrete-time ideal, 384
  - ideal, 318
- LTI systems, 427–39
- LTI systems analysis and, 316
- open-loop, 821
- raised cosine, 629
- rational, Bode plots for, 456–60
- second-order systems, 677–78, 680

Frequency scaling of continuous-time Fourier transform, 308–9

Frequency-selective filter, 231, 236–39, 318, 423
- bandpass, 237–38, 326
- highpass, 237–38, 241–44, 374–75, 387
- ideal, 237–39
  - time-domain properties of, 439–44
- lowpass, 237
- variable center frequency, 325–27

Frequency-shaping filters, 231, 232–36
- differentiating filters, 232–34, 235
- discrete-time LTI filters, 234–36

Frequency shifting property
- continuous-time Fourier series, 206
- continuous-time Fourier transform, 311, 328
- discrete-time Fourier series, 221

  discrete-time Fourier transform, 373–75
Frequency shift keying (FSK), 646
Fundamental components, 187
Fundamental frequency, 17–18
Fundamental period, 12
  continuous-time periodic signal, 186
  discrete-time periodic signal, 211

Gain, 428
  linear feedback systems, 835, 858–66
  tracking system, 829–30
General complex exponentials, 20–21, 24–25
Generalized functions, 36, 136
Gibbs phenomenon, 200–201, 219, 294
Group delay, 430–36

Half-sample delay, 543–45
Hanning window, 422
Harmonically related complex exponentials, 19
Harmonic analyzer, 200
Harmonic components, 187
Heat propagation and diffusion, 180
Higher order holds, 526–27
Highpass filter, 237–38
  frequency response of, 374–75
  ideal bandstop characteristic, 387
  RC, 241–44
Highpass-to-lowpass transformations, 498
Hilbert transform, 351
Hold(s)
  higher order, 526–27
  zero-order, 520–22, 523–26
Homogeneity (scaling) property, 53

Ideal frequency-selective filter, 237–39
  time-domain properties of, 439–44
Idealization, 67
Identity system, 44
Image processing
  differentiating filters for, 232–34, 235
  phase representation and, 425–27
Imaginary part
  complex number, 71
  Fourier series coefficients, 216, 217
Impulse response
  absolutely integrable, 114
  absolutely summable, 113
  associated with group delay, 435–36
  causal LTI system, 112
  continuous-time ideal lowpass filter, 442
  discrete-time ideal lowpass filter, 442
  first-order discrete-time system, 461–62
  ideal lowpass filter, 321–22, 384
  second-order systems, 677–78
    discrete-time systems, 466–68
Impulse train
  continuous-time Fourier series of, 208–10
  continuous-time Fourier transform of, 299–300
  discrete-time Fourier transform of, 371–72, 376
Impulse-train sampling, 516–20, 545–49
Incrementally linear systems, 56
Independent variable, 3–4
  transformation of, 7–14
Infinite impulse response (IIR) filters, 123, 476
Infinite sum formula, 73, 89
Infinite Taylor series, 277
Initial-value theorem
  Laplace transform, 690–91
  z-transform, 773–74
Instantaneous frequency, 613
Integral, convolution, 90–102
  evaluating, 97–102
Integration, in time-domain, 690, 717
Integration property
  continuous-time Fourier series, 206
  continuous-time Fourier transform, 306–8
Integrator, 126–27
Interconnection
  convolution property and analysis of, 386–87
  distributive property of convolution and, 105
  feedback, 43

  LTI systems
    Laplace transform for, 707–8
    system functions for, 707–8
    z-transform for, 784
  parallel, 42–43
Interference, intersymbol, 607–10
Intermediate-frequency (IF) stage, 596
Interpolation, 549–55
  band-limited, 523–24
  linear, 522–23, 526
  reconstruction using, 522–27
Intersymbol interference (ISI), 607–10
Inverse Fourier transform, 284, 288
Inverse Laplace transform, 670–73
Inverse LTI system, 734
Inverse system design, 820–21
Inverse systems, 45–46
Inverse z-transform, 757–63
  examples of, 758–763
Inverted pendulum, 818–19
Invertibility of LTI systems, 109–11
Invertible systems, 45–46

Lacroix, S.F., 180
Lag, phase, 491–92
Lag network, 890
Lagrange, J.L., 178–79, 180
Laguerre polynomials, 738
Laplace, P.S. de, 180
Laplace transform, 654–740
  bilateral, 655, 714
  geometric evaluation of, 674–82
    all-pass systems, 681–82
    first-order systems, 676–77
    second-order systems, 677–81
  inverse, 670–73
  LTI system analysis and characterization using, 693–706
    causality, 693–95, 697
    linearity constant-coefficient differential equations, 698–700
    stability, 695–98
  poles of, 660
  pole-zero plot of, 660–62, 674–82
  properties of, 682–92
    conjugation, 687
    convolution, 687–88
    differentiation in s domain, 688–90
    differentiation in time-domain, 688
    initial- and final-value theorems, 690–91
    integration in time-domain, 690
    linearity, 683–84
    s-domain shifting, 685
    table of, 691–92
    time scaling, 685–87
    time shifting, 684
  region of convergence for, 657–58, 662–70
    left-sided signal, 666, 669
    rational transform, 669
    right-sided signal, 665–66, 669
    two-sided signal, 666–68
  s-plane representation of, 660, 663–64
  as system function, 693, 701–13
    block diagram representations, 708–13
    Butterworth filters, 703–6
    interconnections of LTI systems, 707–8
  transform pairs, 692–93
  unilateral, 714–20
    examples of, 714–16
    properties of, 716–19
    solving differential equations using, 719–20
  zeros of, 660
Lead, phase, 491–92
Lead network, 890
Left-half plane, 666
Left-sided signal, 666
Linear, time-invariant (LTI) systems, 41, 56, 74–176. *See also* Linear constant-coefficient difference equations; Linear constant-coefficient differential equations
  bandwidth of, 352–53
  causal, 46–48, 112–13, 116–27
    block diagram representations for, 124–27, 708–13, 784–89
    described by linear constant-coefficient difference equations, 116, 121–24, 396–99

    described by linear constant-coefficient differential equations, 116, 117–21
    real-part sufficiency property of, 417
  continuous-time, 90–102
    convolution-integral representation of, 94–102
    described by linear constant-coefficient differential equations, 330–33
    representation in terms of impulses, 90–94
  convolution property and, 319–21
  discrete-time, 75–90
    convolution-sum representation of, 77–90
    representation in terms of impulses, 75–77
  feedback. *See* Linear feedback systems
  filtering with, 231, 234–36
  Fourier series and, 226–31
  frequency response of, 316
    magnitude-phase representation of, 427–39
  interconnection of
    Laplace transform for, 707–8
    z-transform for, 784
  inverse, 109–11, 734
  Laplace transform to analyze and characterize, 693–706
    causality, 693–95, 697
    linearity constant-coefficient differential equations, 698–700
    stability, 695–98
  with and without memory, 108–9
  properties of, 103–8
    associative, 107–8
    commutative, 104
    distributive, 104–6
  response to complex exponentials, 182–86, 226–31
  stability for, 113–15
  system functions for interconnections of, 707–8
  unit impulse response of cascade of two, 107
  unit step response of, 115–16
  windowing in design of, 420–21
  z-transform to analyze and characterize, 774–83
    causality, 776–77
    stability, 777–79
Linear constant-coefficient difference equations, 116, 121–24, 396–99, 779–81
  finite impulse response (FIR) system, 122
  infinite impulse response (IIR) system, 123
  natural responses as solutions to, 122
  nonrecursive, 122
  recursive, 122, 123
  unilateral *z*-transform to solve, 795–96
Linear constant-coefficient differential equations, 116, 117–21, 330–33, 698–700
  block diagram representation of, 124–27
  initial rest condition, 119–20
  natural responses as solutions to, 119, 121
  particular and homogeneous solutions to, 118–19
  unilateral Laplace transform to solve, 719–20
Linear feedback systems, 816–908. *See also* Feedback
  closed-loop, 818
  gain and phase margin, 835, 858–66
  Nyquist stability criterion, 846–58
    continuous-time, 850–55
    discrete-time, 856–58
    encirclement property, 847–50
  open-loop, 818
  properties and uses of, 819–20
  root-locus analysis of, 832–46
    angle criterion, 836–40
    end points, 836

    - equation for closed-loop poles, 834–36
    - properties of, 841–46
  - sampled-data, 826–28
- Linear interpolation, 522–23, 526
- Linearity, 53–56
- Linearity property
  - continuous-time Fourier series, 202, 206
  - continuous-time Fourier transform, 301
  - discrete-time Fourier series, 221
  - discrete-time Fourier transform, 373
  - Laplace transform, 683–84
    - unilateral, 717
  - z-transform, 767
    - unilateral, 793
- Log magnitude-phase diagram, 861–62, 864–66
- Log-magnitude plots, 436–39
- Lossless coding, 46
- Lowpass filters, 237
  - frequency response of, 374–75
  - ideal
    - continuous-time, 440, 442, 443
    - discrete-time, 384, 442, 443
    - frequency response of, 318
    - impulse responses of, 321–22, 384, 442
    - step response of, 443
  - *RC*, 239–41
- Lowpass-to-highpass transformations, 498
- LTI. *See* Linear, time-invariant (LTI) systems

- Magnitude of complex number, 71
- Magnitude-phase representation
  - Fourier transform, 423–27
  - frequency response of LTI systems, 427–39
    - group delay, 430–36
    - linear and nonlinear phase, 428–30
    - log-magnitude and Bode plots, 436–39
- Mass-spring-dashpot mechanical system, 824
- Matched filter, 170–72, 275
- Measuring devices, imperfect, 355–56
- Memory, LTI systems with and without, 108–9
- Memoryless systems, 44
- Michelson, Albert, 200
- Modulating signal, 583
- Modulation. *See also* Sinusoidal amplitude modulation
  - angle, 611–13
  - defined, 582
  - discrete-time, 619–23
  - double-sideband (DSB), 598
  - percent, 591
  - phase, 611–13
  - pulse-amplitude, 604–10
    - digital, 610
    - intersymbol interference in, 607–10
  - pulse-code, 610
  - sinusoidal frequency, 583, 611–19
    - narrowband, 613–15
    - periodic square-wave modulating signal, 617–19
    - wideband, 615–17
- Modulation index, 591, 614
- Modulation property, 323. *See also* Multiplication property
- Monge, G., 180
- Moving-average filters, 245–47, 476–82
- Multiplexing
  - defined, 583
  - frequency-division (FDM), 594–97
  - quadrature, 646
  - time-division (TDM), 604, 605
- Multiplication property
  - continuous-time Fourier series, 204, 206
  - continuous-time Fourier transform, 322–27
  - discrete-time Fourier series, 221, 222
  - discrete-time Fourier transform, 388–90
- Multiplicities, 155

- Narrowband frequency modulation, 613–15
- Natural frequency, undamped, 452–53

Natural responses, 119, 121, 122
Negative (degenerative) feedback, 822, 831–32
Networks, lag and lead, 890
Nonanticipative system, 46–47
Noncausal averaging system, 47
Nonideal filters, time-domain and frequency-domain characterization of, 444–47
Nonrecursive filters. *See* Finite impulse response (FIR) filters
Nonrecursive linear constant-coefficient difference equations, 122
Normalized functions, 273
Nyquist frequency, 519
Nyquist plots, 853–58
Nyquist rate, 519, 556
Nyquist stability criterion, 846–58
  continuous-time, 850–55
  discrete-time, 856–58
  encirclement property, 847–50

Odd harmonic continuous-time periodic signal, 262
Odd signals, 13–14
  continuous-time Fourier series, 206
  discrete-time Fourier series, 221
Open-loop system, 818
Open-loop frequency response, 821
Operational amplifiers, 821, 896–97
Orthogonal functions, 273–75
Orthogonal signals, 280–81
Orthonormal functions, 273–75
Orthonormal signals, 280–81
Oscilloscope, sampling, 534
Overdamped system, 453
Oversampling, 529
Overshoot, 454

Parallel-form block diagrams, 712, 713, 787–89
Parallel interconnection, 42–43
Parity check, 610
Parseval's relation
  continuous-time Fourier transform, 312–14
  continuous-time Fourier series, 205, 206
  discrete-time Fourier transform, 380–82
  discrete-time Fourier series, 223
Partial-fraction expansion, 909–20
  continuous-time signals and systems and, 910–16
  discrete-time signals and systems and, 916–20
Passband edge, 445
Passband frequency, 237
Passband ripple, 445
Pendulum, inverted, 818–19
Percent modulation, 591
Period, sampling, 516
Periodic complex exponentials, 16–20, 186
Periodic convolution, 222, 389–90
  continuous-time Fourier series, 206
  discrete-time Fourier series, 221
Periodicity property of discrete-time Fourier transform, 373
Periodic signals, 11–13. *See also* Fourier transform; Fourier series; Periodic complex exponentials; Sinusoidal signals
  continuous-time, 12
  discrete-time, 12
  power of, 312
Periodic square wave
  continuous-time, 193–95, 200–201, 209, 218–19, 285–86, 297–98
  discrete-time, 219–20, 224
Periodic train of impulses
  continuous-time Fourier series of, 208–10
  continuous-time Fourier transform of, 299–300
  discrete-time Fourier transform of, 371–72, 376
Periodic signal, odd-harmonic, 262
Phase (angle) of complex number, 71
Phase lag, 491–92
Phase lead, 491–92
Phase margin in linear feedback systems, 858–66

Phase modulation, 611–13
Phase reversal, 532
Phase shift, 428, 636
Plant, 828
Polar form for complex number, 71
Poles
  closed-loop, 834–36
  Laplace transform, 660
Pole-zero plot(s)
  all-pass systems, 682
  first-order systems, 676–77, 763–65
  Laplace transforms, 660–62, 674–82
  second-order systems, 678–80, 765–67
  z-transform, 763–67
Polynomials, Laguerre, 738
Population dynamics, feedback model of, 824–26
Positive (regenerative) feedback, 831–32
Power of signals, 5–7, 312
Power-series expansion method, 761–63
Principal-phase function, 433, 434
Proportional feedback system, 823–24
Proportional (P) control, 905
Proportional-plus-derivative feedback, 824
Proportional-plus-integral (PI) control, 905
Proportional-plus-integral-plus-differential (PID) control, 905
Pulse-amplitude modulation, 604–10
  digital, 610
  intersymbol interference in, 607–10
Pulse-code modulation, 610
Pulse-train carrier, 601–4, 605

Quadrature distortion, 636
Quadrature multiplexing, 646
Quality of circuit, 456

Raised cosine frequency response, 629
Rational frequency responses, Bode plots for, 456–60
RC highpass filter, 241–44
RC lowpass filter, 239–41
Real exponential signals, 15, 22, 23
Real part
  complex number, 71
  Fourier series coefficients, 216, 217
Real-part sufficiency, 350, 417
Rectangular (Cartesian) form for complex number, 71
Rectangular pulse
  continuous-time, 293–94
  discrete-time, 365–66
Rectangular window, 420
Recursive (infinite impulse response) filters, 123, 476
Recursive linear constant-coefficient difference equations, 122, 123
Regenerative (positive) feedback, 831–32
Region of convergence
  Laplace transform, 657–58, 662–70
    left-sided signal, 666, 669
    rational transform, 669
    right-sided signal, 665–66, 669
    two-sided signal, 666–68
  z-transform, 743–44, 748–57
    bounded by poles or infinity, 756
    centered about origin, 748–49
    finite-duration sequence, 749–50
    left-sided sequence, 751, 752
    right-sided sequence, 750–51, 752
    two-sided sequence, 751–53
Right-half plane, 666
Right-sided signal, 665–66
Ringing, 443, 454
Rise time, 352–53, 444
Root-locus analysis, 832–46
  angle criterion, 836–40
  end points, 836
  equation for closed-loop poles, 834–36
  properties of, 841–46
Running sum of discrete-time Fourier series, 221

Sampled-data feedback systems, 826–28
Sampling, 29, 39, 514–81
  aliasing, 527–34
  bandpass signal, 564–65
  continuous-time signals, 4

  discrete-time processing of continuous-time signals, 534–45
    digital differentiator, 541–43
    half-sample delay, 543–45
  discrete-time signals, 545–55
    decimation and interpolation, 549–55
    impulse train, 545–49
  impulse-train, 516–20, 545–49
  reconstruction using interpolation, 522–27
  with zero-order hold, 520–22, 523–26
Sampling frequency, 516
Sampling function, 516
Sampling oscilloscope, 534
Sampling period, 516
Sampling property of continuous-time impulse, 35
Sampling theorem, 514–15, 518
Scaling (homogeneity) property, 53
Scaling in z domain, 768–69, 793
Scrambler, speech, 633–34
Second harmonic components, 187
Second-order continuous-time systems, 451–56
Second-order discrete-time systems, 465–72
Series (cascade) interconnection, 42
Shifting property
  s-domain, 717
  unilateral z transform, 794–95
Sifting property
  continuous-time impulse, 90–92
  discrete-time unit impulse, 77
Signal(s), 1–38. *See also* Periodic signals
  continuous-time and discrete-time, 1–7
    energy and power of, 5–7
    examples and mathematical representation of, 1–5
    sampling of, 4, 545–55
  exponential and sinusoidal, 14–30, 186
    continuous-time complex, 15–21
    discrete-time complex, 21–30
    real, 22, 23
  transformation of independent variable, 7–14
    examples of, 8–11
  unit impulse and unit step functions, 30–38
    continuous-time, 32–38
    discrete-time, 30–32
Sinc functions, 295
Single-sideband sinusoidal amplitude modulation, 597–601
Singularity functions, 36, 67, 127–36. *See also* Unit impulse
Sinusoidal amplitude modulation, 583–87
  complex exponential carrier, 583–85
  demodulation for, 587–94
    asynchronous, 590–94
    synchronous, 587–90
  discrete-time, 619–23
  frequency-division multiplexing (FDM) using, 594–97
  single-sideband, 597–601
  with sinusoidal carrier, 585–87
Sinusoidal frequency modulation, 583, 611–19
  narrowband, 613–15
  periodic square-wave modulating signal, 617–19
  wideband, 615–17
Sinusoidal signals, 14–30, 186
  continuous-time complex, 15–21
  discrete-time, 24
  discrete-time complex, 21–30
  real, 22, 23
Sinusoids, damped, 21
Sliding, 85
Spectral coefficient. *See* Fourier series coefficients
Spectrum, 288, 361. *See also* Fourier transform
  energy-density, 312, 349, 381
Speech scrambler, 633–34
Square wave, periodic
  continuous-time, 193–95, 200–201, 209, 218–19, 285–86, 297–98
  discrete-time, 219–20, 224

Stability, 48–50
  feedback systems, 858–66
  LTI systems, 113–15, 695–98, 777–79
Stabilization with feedback, 823–26
Step-invariant transformation, 828, 908
Step response, 115–16
  automotive suspension system, 476
  continuous-time ideal lowpass filter, 443
  discrete-time ideal lowpass filter, 443
  first-order discrete-time system, 461, 463
  second-order discrete-time systems, 466, 468
Stopband edge, 445
Stopband frequency, 237
Stopband ripple, 445
Stroboscopic effect, 533–34
Sufficiency, real-part, 350, 417
Summer, 44
Superposition integral. *See* Convolution integral
Superposition property, 53
Superposition sum. See Convolution sum
Suspension system, analysis of, 473–76
Symmetry, conjugate, 204–5, 206, 221, 303–6, 375
Synthesis equation
  continuous-time Fourier series, 191
  continuous-time Fourier transform, 288, 300, 314
  discrete-time Fourier series, 213
  discrete-time Fourier transform, 361, 390
System function(s), 227
  closed-loop, 820
  feedback path, 820
  forward path, 820
  interconnections of LTI systems, 707–8
  Laplace transform as, 693, 701–13
    block diagram representations, 708–13
    Butterworth filters, 703–6
    interconnections of LTI systems, 707–8
  z-transform as, 781–89
    block diagram representations, 784–89
    interconnection of LTI systems, 784
Systems, 38–56
  continuous-time and discrete-time, 38–43
    examples of, 39–41
    first-order, 448–51, 461–65
    interconnections of, 41–43
    second-order, 451–56, 465–72
  properties of, 44–56
    causality, 46–48
    invertible and inverse, 45–46
    linearity, 53–56
    memory and memoryless, 44–45
    stability, 48–50
    time-invariance, 50–53

Taylor series, infinite, 277
Telescope, angular position of, 816–18
Time, 3–4
Time advance property of unilateral z-transform, 793
Time constants, 448
  dominant, 500–501
Time delay property of unilateral z-transform, 793
Time-division multiplexing (TDM), 604, 605
Time-domain and frequency-domain characterization, 423–513
  continuous-time systems, 448–60
    Bode plots for rational frequency responses and, 456–60
    first-order, 448–51
    second-order, 451–56
  discrete-time systems, 461–72
    first-order, 461–65
    second-order, 465–72
  examples of, 472–82
    automobile suspension system, 473–76
    discrete-time nonrecursive filter, 476–82
  magnitude-phase representation
    Fourier transform, 423–27

    - frequency response of LTI systems, 427–39
  - nonideal filters, 444–47
- Time expansion property
  - discrete-time Fourier series, 221
  - discrete-time Fourier transform, 377–80
  - z-transform, 769–70
    - unilateral, 793
- Time invariance, 50–53, 77
- Time reversal property, 8
  - continuous-time Fourier series, 203, 206
  - continuous-time Fourier transform, 309
  - discrete-time Fourier series, 221
  - discrete-time Fourier transform, 376–77
  - Laplace transform, 687
  - z-transform, 769
- Time scaling property, 8, 9, 10
  - continuous-time Fourier series, 204, 206
  - continuous-time Fourier transform, 308–9
  - discrete-time Fourier series, 221
  - discrete-time Fourier transform, 377–80
  - Laplace transform, 685–87
    - unilateral, 717
  - z-transform, 769–70
    - unilateral, 793
- Time shifting property, 8–9, 10
  - continuous-time Fourier series, 202–3, 206
  - continuous-time Fourier transform, 301–3
  - discrete-time Fourier series, 221
  - discrete-time Fourier transform, 373–75
  - Laplace transform, 684
  - z-transform, 767–68
- Time window, 420
- Tracking systems, 828–30
- Transfer function. *See* System function(s)
- Transformation
  - bilinear, 814–15
  - highpass-to-lowpass, 498
  - of independent variable, 7–14
  - lowpass-to-highpass, 498
  - step-invariant, 828, 908
- Transition band, 445
- Transmodulation (transmultiplexing), discrete-time, 623, 624
- Triangular (Bartlett) window, 421
- Trigonometric series, 178–79
- Two-sided signal, 666–68
- Type l feedback system, 906

- Undamped natural frequency, 452–53
- Undamped system, 454
- Underdamped systems, 467
- Undersampling, 527–34
- Unilateral Laplace transform, 714–20
  - examples of, 714–16
  - properties of, 716–19
  - solving differential equations using, 719–20
- Unilateral z-transform, 742, 789–96
  - examples of, 779–92
  - properties of, 792–95
  - solving difference equations using, 795–96
- Unit circle, 743
- Unit delay, 125
- Unit doublets, 132–36, 172
- Unit impulse, 30–38
  - cascade of two LTI systems and, 107
  - continuous-time, 32–38, 127–36, 292
    - defining through convolution, 131–32
    - as idealized short pulse, 128–31
    - sifting property of, 90–92
  - discrete-time, 30–32, 367
    - sifting property of, 77
  - representation of continuous-time LTI systems in terms of, 90–94
  - representation of discrete-time LTI systems in terms of, 75–77
- Unit impulse response, discrete-time, 77–90. *See also* Impulse response

Unit ramp function, 135
Unit step functions, 30–38
  continuous-time, 32–38
  discrete-time, 30–32
Unit step response of LTI systems, 115–16
Unstable systems, feedback to stabilize, 823–26
Unwrapped phase, 433, 434
Upsampling, 552, 553

Variable, independent, 3–4
  transformation of, 7–14

Walsh functions, 170, 276
Wave, square. *See* Periodic square wave
Wavelengths, 181
Weighted average, 245
Wideband frequency modulation, 615–17
Windowing, 420–22

Zero-force equalizer, 649
Zero-input response, 55–56, 720
Zero-order hold, 520–22, 523–26
Zeros of Laplace transform, 660
Zero-state response, 720
z-transform, 741–815
  bilateral, 742, 789
  defined, 742
  discrete-time Fourier transform and, 743
  geometric evaluation of, 763–67
    first-order systems, 763–65
    second-order systems, 765–67
  inverse, 757–63
    examples of, 79–92
LTI systems analysis and characterization using, 774–83
  causality, 776–77
  linear constant-coefficient difference equations, 779–81
  stability, 777–79
pole-zero plot, 763–67
properties of, 767–74
  conjugation, 770
  convolution, 770–72
  differentiation in *z*-domain, 772
  initial-value theorem, 773–74
  linearity, 767
  scaling in *z*-domain, 768–69
  table of, 775
  time expansion, 769–70
  time reversal, 769
  time shifting, 767–68
region of convergence for, 743–44, 748–57
  bounded by poles or infinity, 756
  centered about origin, 748–49
  finite-duration sequence, 749–50
  left-sided sequence, 751, 752
  right-sided sequence, 750–51, 752
  two-sided sequence, 751–53
system functions, 781–89
  block diagram representations, 784–89
  for interconnection of LTI systems, 784
unilateral, 742, 789–96
  examples of, 789–796
  properties of, 792–95
  solving difference equations using, 795–96
z-transform pairs, 774, 776